

Db	61	ATGAGCGTCGTATTCGGACCGAGACCAAAAACAAAGAAATTCGCGGCTGCCATCACCCG	120
Qy	121	GCCGCGCTCGCGAATCTAACCCCTCGTGGCATAGAGTCTCTATCCAGGCAAGTCCGGA	180
Db	121	GCCGCGCTCGCGAATCTAACCCCTCGTGGCATAGAGTCTCTATCCAGGCAAGTCCGGA	180
Qy	181	GAGGCGCTCGGCTATCACCGACGGGAAATTCGAAGGCGCGAGGCGCACTGGTCGGAC	240
Db	181	GAGGCGCTCGGCTATCACCGACGGGAAATTCGAAGGCGCGAGGCGCACTGGTCGGAC	240
Qy	241	GCCGACCAAGTGTGTGGCCGACGCTGATTTATGTCTCAAGTCAAAAGAACGATAGCGCG	300
Db	241	GCCGACCAAGTGTGTGGCCGACGCTGATTTATGTCTCAAGTCAAAAGAACGATAGCGCG	300
Qy	301	GAATACGGCGCGCTCGACACGGGAGATCTTGTACGTTCTTCGATTTGGCCGCGCA	360
Db	301	GAATACGGCGCGCTCGACACGGGAGATCTTGTACGTTCTTCGATTTGGCCGCGCA	360
Qy	361	CGTGTTCACACGATGGTGTGTGGATTCGCGCACACGTCATTTGCTACGACGACCTC	420
Db	361	CGTGTTCACACGATGGTGTGTGGATTCGCGCACACGTCATTTGCTACGACGACCTC	420
Qy	421	CAGACCGCCGACGGCGCACTAACCCCTGTGTTGCCCGCATACGAAAGTGGCGGTGACTC	480
Db	421	CAGACCGCCGACGGCGCACTAACCCCTGTGTTGCCCGCATAGGAAAGTGGCGGTGACTC	480
Qy	481	GCCGCGCAGGTGTGGCGCTTACCACTGATCGAAACCAAGGGGGCGCGGTGCTGATG	540
Db	481	GCCGCGCAGGTGTGGCGCTTACCACTGATCGAAACCAAGGGGGCGCGGTGCTGATG	540
Qy	541	GCGCGGCGTCCCGCGCTCGAACCGGCGACGCTGTGTATCGGCGCGCGCACCGCGCGC	600
Db	541	GCGCGGCGTCCCGCGCTCGAACCGGCGACGCTGTGTATCGGCGCGCGCACCGCGCGC	600
Qy	601	TACAAACGACGCCCATCTGCGCAACGGGATGGGCGCACCGTTACGTTCTAGACATCAAC	660
Db	601	TACAAACGACGCCCATCTGCGCAACGGGATGGGCGCACCGTTACGTTCTAGACATCAAC	660
Qy	661	ATCGACAACCTTCGGCAACTCGAAGCGCGAATTTGCGGCGCGATCGACACTGGTACTCA	720
Db	661	ATCGACAACCTTCGGCAACTCGAAGCGCGAATTTGCGGCGCGATCGACACTGGTACTCA	720
Qy	721	TCGCGCTACGAGCTCGAGGGTGCCTGAACAGTGCCGACCTGGTGTGTGGGCGTCTCTG	780
Db	721	TCGCGCTACGAGCTCGAGGGTGCCTGAACAGTGCCGACCTGGTGTGTGGGCGTCTCTG	780
Qy	781	GTCGACGGCGCAGGAGCAACCAATTGTCGAAATTCACCTTGTCCGATATGAACA	840
Db	781	GTCGACGGCGCAGGAGCAACCAATTGTCGAAATTCACCTTGTCCGATATGAACA	840
Qy	841	GGTGTGCTACGTGTGATATATAGCCATAGACAGGCGGCGTGTTCGAAGGCTCACGACG	900
Db	841	GGTGTGCTACGTGTGATATATAGCCATAGACAGGCGGCGTGTTCGAAGGCTCACGACG	900
Qy	901	ACCACTACGACACACCGAGCTTGCCGCTGACAGACAGCTGTTTACGTGCTGGCAAC	960
Db	901	ACCACTACGACACACCGAGCTTGCCGCTGACAGACAGCTGTTTACGTGCTGGCAAC	960
Qy	961	ATGCCCGCTCGGTGCGCAAGACGTGACCTAGCGGCTGACCAACGACGAGATGCGGTAT	1020
Db	961	ATGCCCGCTCGGTGCGCAAGACGTGACCTAGCGGCTGACCAACCGAGATGCGGTAT	1020
Qy	1021	GTGCTCGAGCTTGGCGACATATGCTGTGGCGAGCGCGTGTGCAATCCGACACTAGCG	1080
Db	1021	GTGCTCGAGCTTGGCGACATATGCTGTGGCGAGCGCGTGTGCAATCCGACACTAGCG	1080
Qy	1081	AAAGGCTTTTGCAGCGACGAAGGCGGCTTACTGTCCGAAGGCGTGGCCACCACTCGGGG	1140
Db	1081	AAAGGCTTTTGCAGCGACGAAGGCGGCTTACTGTCCGAAGGCGTGGCCACCACTCGGGG	1140
Qy	1141	GTGCGCTTACCGAGCCCGGACGAGGTGTGTGGCGTGAATTCGGCGCGCTGTATACGCGAG	1200
Db	1141	GTGCGCTTACCGAGCCCGGACGAGGTGTGTGGCGTGAATTCGGCGCGCTGTATACGCGAG	1200

[illegible]

QY 601 TACAGCAGACCCGATCGGCAAGCGGATGGGCGACCGTTACGGTTCTAGACATCAAC 660
 DB 601 TACAGCAGACCCGATCGGCAAGCGGATGGGCGACCGTTACGGTTCTAGACATCAAC 660
 QY 661 ATGACAACTTGGGCACTGACGCGGATTTGCGGCGGATTCACACTGCTACTCA 720
 DB 661 ATGACAACTTGGGCACTGACGCGGATTTGCGGCGGATTCACACTGCTACTCA 720
 QY 721 TCGGCGCTAGAGCTCGAGGGTGGCGTCAAAAGTGGCGAGCT...TGATGGGGCGGCTG 780
 DB 721 TCGGCGCTAGAGCTCGAGGGTGGCGTCAAAAGTGGCGAGCT...TGATGGGGCGGCTG 780
 QY 781 GTGCGAGCGCCCAAGCGACCAATTAATGCTCGAATTCATTTGCGGCTATGAACCA 840
 DB 781 GTGCGAGCGCCCAAGCGACCAATTAATGCTCGAATTCATTTGCGGCTATGAACCA 840
 QY 841 GGTGCGGCTAGTGTGATATAGCATGACAGGCGGCTGTTTCAAGGCTACGACCG 900
 DB 841 GGTGCGGCTAGTGTGATATAGCATGACAGGCGGCTGTTTCAAGGCTACGACCG 900
 QY 901 ACCGACCTAGACACCGGAGGTTGCGGCTGACAGACGCGTTTACTGCGTGGGCAAC 960
 DB 901 ACCGACCTAGACACCGGAGGTTGCGGCTGACAGACGCGTTTACTGCGTGGGCAAC 960
 QY 961 ATGCGCGCTCGGTGCGGAGAGCTGACCTACGCGCTGACCAAGCGGACGATCGCTAT 1020
 DB 961 ATGCGCGCTCGGTGCGGAGAGCTGACCTACGCGCTGACCAAGCGGACGATCGCTAT 1020
 QY 1021 GTGCTGAGCTTGGCGGACGATGGTGGCGGCGGCGGCTGCGATCGGCACTAGCC 1080
 DB 1021 GTGCTGAGCTTGGCGGACGATGGTGGCGGCGGCGGCTGCGATCGGCACTAGCC 1080
 QY 1081 AAAGCTTTTGAAGCAGCAAGGCGGCTTACTGTCGCAAGGCGTGGCGACGACTGGG 1140
 DB 1081 AAAGCTTTTGAAGCAGCAAGGCGGCTTACTGTCGCAAGGCGTGGCGACGACTGGG 1140
 QY 1141 GTGCGCTTACACGAGCCCGGCGGCTGCTGCTGACTCTGCGCGCTGTTACGCGGAG 1200
 DB 1141 GTGCGCTTACACGAGCCCGGCGGCTGCTGCTGACTCTGCGCGCTGTTACGCGGAG 1200
 QY 1201 CACACGTGGGAGTAAAGGAGGATGA 1228
 DB 1201 CACACGTGGGAGTAAAGGAGGATGA 1228
 RESULT 4
 LOCUS A89747 1237 bp DNA PAT 22-JAN-2000
 DEFINITION Sequence 4 from Patent W09832862.
 ACCESSION A89747
 VERSION A89747.1 GI:6738281
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1237)
 AUTHORS Flohe, L. and Singh, M.
 TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
 JOURNAL Patent: WO 9832862-A 30-JUL-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source Location/Qualifiers
 1..1237
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 236 a 394 c 386 g 221 t
 ORIGIN
 Query Match 100.0%; Score 1228; DB 5; Length 1237;
 Best Local Similarity 100.0%; Pred. No. 4.2e-169;
 Matches 1228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATCTTCAGATTAATCGAACTTTCTTCACTACTGAAGCTACAGTATCGAGGCGGTAAATC 60

DB 1 ATCTTCAGATTAATCGAACTTTCTTCACTACTGAAGCTACAGTATCGAGGCGGTAAATC 60
 QY 61 ATGCGGTGGTATTTCCAGCAGACAAAAACAAGATTCGGGGTCCATCACCCG 120
 DB 61 ATGCGGTGGTATTTCCAGCAGACAAAAACAAGATTCGGGGTCCATCACCCG 120
 QY 121 GCGGCGTGGGAACTAACCGGTGGGCTATGAGGTGCTCATCGAGCAGAGTCCGGA 180
 DB 121 GCGGCGTGGGAACTAACCGGTGGGCTATGAGGTGCTCATCGAGCAGAGTCCGGA 180
 QY 181 GAGGCTCGGCTATACCGCAGCGGATTTCAAGGGCGGCGCAACTGGTGGCAC 240
 DB 181 GAGGCTCGGCTATACCGCAGCGGATTTCAAGGGCGGCGCAACTGGTGGCAC 240
 QY 241 GCGGACAGAGTGGGCGCAGCTGATTTATGCTCAAGGTCAAAAGAACCTATACGGG 300
 DB 241 GCGGACAGAGTGGGCGCAGCTGATTTATGCTCAAGGTCAAAAGAACCTATACGGG 300
 QY 301 GAATACGGCGCGCTGCGAGACGGGAGATCTTTCACCTTCTGCAATTTGCGCGTCA 360
 DB 301 GAATACGGCGCGCTGCGAGACGGGAGATCTTTCACCTTCTGCAATTTGCGCGTCA 360
 QY 361 CGTGTTCACAGGATCGCTGTTGGATTCCGACACCTCAATTGCTTACGAGACGCTC 420
 DB 361 CGTGTTCACAGGATCGCTGTTGGATTCCGACACCTCAATTGCTTACGAGACGCTC 420
 QY 421 CAGACCGCGGAGGCGCATACCCCTGCTTGGCGCGGATAGGAAAGTGGCGGTGAGTC 480
 DB 421 CAGACCGCGGAGGCGCATACCCCTGCTTGGCGCGGATAGGAAAGTGGCGGTGAGTC 480
 QY 481 GCGCGCCAGGTTGGCGCTTACACCTGATCGCAACCAAGGGGCGCGGTGTGTGATG 540
 DB 481 GCGCGCCAGGTTGGCGCTTACACCTGATCGCAACCAAGGGGCGCGGTGTGTGATG 540
 QY 541 GCGGCGGTGGCGGCTCGAACCGGCGACGCTGTGTGATGAGGGCGCGGACCGCGG 600
 DB 541 GCGGCGGTGGCGGCTCGAACCGGCGACGCTGTGTGATGAGGGCGCGGACCGCGG 600
 QY 601 TACAGCAGACCCGATCGGCAAGGCGGATGGGCGGCGGTTACGCTTACATCAAC 660
 DB 601 TACAGCAGACCCGATCGGCAAGGCGGATGGGCGGCGGTTACGCTTACATCAAC 660
 QY 661 ATGACAACTTGGGCACTGACGCGGATTTGCGGCGGATTCACACTGCTACTCA 720
 DB 661 ATGACAACTTGGGCACTGACGCGGATTTGCGGCGGATTCACACTGCTACTCA 720
 QY 721 TCGGCGCTAGAGCTCGAGGGTGGCGTCAAAAGTGGCGAGCT...TGATGGGGCGGCTG 780
 DB 721 TCGGCGCTAGAGCTCGAGGGTGGCGTCAAAAGTGGCGAGCT...TGATGGGGCGGCTG 780
 QY 781 GTGCGAGCGCCCAAGCGACCAATTAATGCTCGAATTCATTTGCGGCTATGAACCA 840
 DB 781 GTGCGAGCGCCCAAGCGACCAATTAATGCTCGAATTCATTTGCGGCTATGAACCA 840
 QY 841 GGTGCGGCTAGTGTGATATAGCATGACAGGCGGCTGTTTCAAGGCTACGACCG 900
 DB 841 GGTGCGGCTAGTGTGATATAGCATGACAGGCGGCTGTTTCAAGGCTACGACCG 900
 QY 901 ACCGACCTAGACACCGGAGGTTGCGGCTGACAGACGCGTTTACTGCGTGGGCAAC 960
 DB 901 ACCGACCTAGACACCGGAGGTTGCGGCTGACAGACGCGTTTACTGCGTGGGCAAC 960
 QY 961 ATGCGCGCTCGGTGCGGAGAGCTGACCTACGCGCTGACCAAGCGGACGATCGCTAT 1020
 DB 961 ATGCGCGCTCGGTGCGGAGAGCTGACCTACGCGCTGACCAAGCGGACGATCGCTAT 1020
 QY 1021 GTGCTGAGCTTGGCGGACGATGGTGGCGGCGGCGGCTGCGATCGGCACTAGCC 1080
 DB 1021 GTGCTGAGCTTGGCGGACGATGGTGGCGGCGGCGGCTGCGATCGGCACTAGCC 1080
 QY 1081 AAAGCTTTTGAAGCAGCAAGGCGGCTTACTGTCGCAAGGCGTGGCGACGACTGGG 1140

Db	1081	AAAGTCTTTGAGCAGCAGGAAGGGCGTTACTGTGTCCGAACGGGATGGCCACGACCACTGGGG	1140
0y	1141	GTGCCGTTACCGAGCAGCCGACGCTGCTGGCTATCTTCGGCCGCTGTTACGCCGAG	1200
Db	1141	GTGGCGTTACCGAGCAGCCGACGCTGCTGGCTATCTTCGGCCGCTGTTACGCCGAG	1200
0y	1201	CACAGCTGGGAGTAAGGAAGCATGA	1228
Db	1201	CACAGCTGGGAGTAAGGAAGCATGA	1228
RESULT	5		
MTV002	56414 bp	DNA	BC1 17-JUN-1998
LOCUS			
DEFINITION	Mycobacterium tuberculosis H37Rv complete genome; segment 122/162.		
ACCESSION	AL008967	AL133456	
VERSION	AL008967.1	GI:3261491	
KEYWORDS			
SOURCE			
ORGANISM	Mycobacterium tuberculosis.		
	Mycobacterium tuberculosis		
	Bacteria: Firmicutes: Actinobacteria: Actinobacteridae;		
	Actinomycetales: Corynebacteriinae; Mycobacteriaceae;		
	Mycobacterium.		
REFERENCE	1 (bases 1 to 56414)		
AUTHORS	Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry III, C.E., Tekle, A.F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Felwell, T., Gentles, S., Hamlin, N., Holtroyd, S., Hornsby, T., Jagsels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, S., Osborne, J., Quail, M.A., Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skelton, S., Squares, S., Squires, R., Sulston, J.E., Taylor, K., Whitehead, S. and Barrrell, B.G.		
TITLE	Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence		
JOURNAL	Nature 393 (6685), 537-544 (1998)		
MEDLINE	98295987		
REMARK	Erratum: [[published erratum appears in Nature 1998 Nov 12;396(6707):1901]]		
REFERENCE	2 (bases 1 to 56414)		
AUTHORS	Parkhill, J.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk		
COMMENT	On Jun 27, 1998 this sequence version replaced gi:2624256.		
Notes:	Details of M. tuberculosis sequencing at the Sanger Centre are available on the World Wide Web. (URL: http://www.sanger.ac.uk/Projects/M.tuberculosis/) CDS have been renumbered from the original cosmid submissions but the old gene designations are in brackets after the new gene numbers. Gene prediction was based on a Hidden Markov Model of TB genes implemented in Tbpase (Krogh) supplemented with visual inspection of positional base preference in codons, especially where there is an increase in the observed/expected third position G + C. CAUTION: In some cases we may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gta, or ttg) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.		
FEATURES			
source	Location/Qualifiers		
	1..56414		
	/organism="Mycobacterium tuberculosis"		
	/strain="H37Rv"		
	/db_xref="taxon:1773"		
	<1..>56414		
	/note="fragment designated v002. Does not represent a physical clone"		
	<1..233		

```

gene
    organism="Mycobacterium tuberculosis"
    /strain="H37Rv"
    /db_xref="taxon:1773"
    /clone="Y154"
    complement(3..527)
    /gene="recX"
    complement(3..527)
    /note="RV2736c, (MTV002.01c), len: 174 aa, recx, similar to eg. RRCX_PSEAE_P37860 regulatory protein recx from Pseudomonas aeruginosa (153 aa), fasta scores: 161 z-score: 257.2 E(): 3.6e-07, 30.7% identity in 137 aa overlap. Overlaps and extends CDS from overlapping cosmid MCV154_16c"
    /codon_start=1
    /transl_table=11
    /product="recX"
    /protein_id="CAI15532.1"
    /db_xref="GI:2624257"
    /db_xref="SPTREMBL:O33280"
    /translation="MTVCPEPSTSEEEQARALCURLTARSRTIAELAQLAKRGT PEDIGNRVLDRIALAGLVDDTDFAEOMVSRRANNAKSKRALAELHAKGVDVVITTT VLGIDGAGEGRRAEKVLVARLRREVLLDGDTEARSRSLVAMLRKGYQLACEPCIVIALAESERRRRV"
    complement(93..2855)
    /gene="recA"
    complement(93..2855)
    /gene="recA"
    /note="RV2737c, (MTV002.02c), len: 790 aa, recA, identical to RECA_MYCTU_P26345 recA protein (790 aa). Contains self-splicing protein element (intron) from 2294 to 969 (c), similar to intron II from TR:E33217 (EMBL:Y11030) DNA-directedDNA polymerase (EC 2.7.7.7) from Thermococcus sp. (1829 aa), fasta scores; opt: 81 z-score 233.2 E(): 6e-05, 24.6%identity in 183 aa overlap. Contains PS00017 ATP/GTP-binding site motif A (P-loop), PS00321 recA signature, and PS00881protein splicing signature. See Davis et al, (1992) Cell11(2):201-210"
    /codon_start=1
    /transl_table=11
    /product="recA"
    /protein_id="CAI15533.1"
    /db_xref="GI:2624258"
    /db_xref="SWISS-PROT:P26345"
    /translation="MTQPDPRKKALELAQAIEKSYKGSVMRLGEAROPISVIPTD SIADLVAGICGLIPRGRIETIGPSSSKTTVAALHANNAAGVARTIDEHALDD DDIARKGVLPDLSLVSPDPDEQALEIDMLIRSGALDIYDVSAALYPRALEEGED GDSHVGLQNALNSQALKRMKTGALNNGSTALF INQLDKIGVFSGPETTGAKLKL YASVRDRRVERVELTKDGTNAVGNTRVNVKNKLAGSTIIPPVTGTHRLIEDVDD RKPIPVAAAKDGTLHARPVVMFDGGTRVIGIRIGGAIWATPDHVKLELYCMQRK AGELRRDQVAORRPREDFGDSAP IPADHALLLYLILRDGMWGGKPIEFINVQAG IWDYTRIADTVNLFLGLFEPSDWNSREQGALRVYTTSQALHQIHMLIFSGW WFEPIRDPIPKRPSTLVNGRIQSRRQVEVRISGMONTVAIESVPIMGPGGAILTI AIPATRGDSRGQATVLAEMTDAVLDELRYSEIVTQEAANAIGVASDPGSGMKQI LGSRSRDGVQVALDALDKFTLHDMLAEELRYSEIVTQEAANAIGVASDPGSGMKQI VAEEVYNHNSPPFOKAEPDIIYGKISRCSLLDMGVDOGILRKSGAFVTEGSQLI QGRKANENILEVNADVADIKEKKIKKIGICAVVTDPSNDGVLPAPVDF"
    complement(790..2112)
    /gene="recA"
    /note="recA intein"
    complement(790..807)
    /gene="recA"
    /note="PS00881 Protein splicing signature"
    complement(2197..2223)
    /gene="recA"
    /note="PS00321 recA signature"
    complement(2644..2667)
    /gene="recA"
    /note="PS00017 ATP/GTP-binding site motif A (P-loop)"
    complement(2873..2877)
    /note="possible RBS upstream of RV2737c"
    complement(3247..3453)

```

CDS
 /gene="RV2738c"
 complement(3247..3453)
 /note="RV2738c"
 N-terminus is highly similar to the N-terminus of the
 upstream ORFMY002.07c (78.4% identity in 37 aa overlap);
 also similar to A020958|SC4H8.5 Streptomyces coelicolor
 cosmid 4H8 (64aa) opt: 185 z-score: 283.5 E(): 2.9e-08;
 39.7% identity in 63 aa overlap
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein RV2738c"
 /protein_id="CA15534.1"
 /db_xref="GI:2624260"
 /db_xref="SPTREMBL:O33281"
 /translation="MLAGVRLTEPHERALHFGAAYGSSVLDHVLTFDGRSAQAI
 EDGVEPRDVMRALCADFDVPHDRW"
 complement(3464..4630)
 /gene="RV2739c"
 complement(3464..4630)
 /note="RV2739c" (MTV002.04c), len: 388 aa. Probable
 transferase. Similar to eg. TR:Q51560 (EMBL:U28170)
 pseudomonas aeruginosa thiamosyl transferase (426 aa);
 fasta scores: opt: 178 z-score: 226.3 E(): 1.9e-05, 25.9%
 identity in 425 aa overlap. Equivalent to Mycobacterium
 leprae protein MCB33.02c (392 aa); fasta scores
 91294723|MCB33.2 Mycobacterium leprae cc nid B33 opt:
 2112 z-score: 2364.5 E(): 0; 80.9% identity in 388 aa
 overlap
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein RV2739c"
 /protein_id="CA15535.1"
 /db_xref="GI:2624261"
 /db_xref="SPTREMBL:O33282"
 /translation="MRVAVVAGPDPGSHSPALALCORPRAADPTLTFTGVEMLEAR
 AGIDAVEIDGLATDRDLDAARIRHRAAOMATLNRRLALPELVSVVITACCG
 MAEELIGIPEWELNPHPLPLPSKGLPPIGSLAGTGRGLRDTMTALGSRWRG
 LRORAVAREIGLIPARDPGLRLIATLPALEVPDPAVAVVGLPHFPPTRVLA
 IPAGPVVVAEPSTALGTAGLLEVALQSITPGTVPVSGSLVSLSGADLVPPV
 AVAGSGSOELLTRADLVICGGHGMVKTLLAGPAMVAVVGGGQMFARVROGS
 AVLIRPADALVAIVNEVLSPPRFEARAAASVAADAPVRVCHALALAC"
 4674..5123
 /gene="RV2740c"
 4674..5123
 /note="RV2740c" (MTV002.05), len: 149 aa. unknown, similar
 to M. leprae hypothetical protein TR:Q49850 (EMBL:294723)
 MCB33.03 (8235..F3_140) (178 aa) fasta scores: opt: 498
 z-score: 538.9 E(): 7.3e-23, 51.6% identity in 161 aa
 overlap
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein RV2740c"
 /protein_id="CA15536.1"
 /db_xref="GI:2624262"
 /db_xref="SPTREMBL:O33283"
 /translation="MAELTETSPETPETTEAIRAVEAFNLALQNEDEFTVDAACDL
 VYENYFSIRIGRRTATLIRMOGRVGEVAVIRHIGAGAVLTERDALIIGPLRV
 OFMVGVFEVDGGRITLMRDYVDYDMKGLRGLIALVAVPSLKATL"
 5355..6932
 /gene="PE_PGRS"
 5355..6932
 /note="PE_PGRS" (MTV002.06), len: 525 aa; Member of M.
 tuberculosis PE_PGRS subfamily, similar to many eg.

Query Match 100.0% Score 1228: DB 1: Length 56414;
 Best Local Similarity 100.0% Pred. No. 1.8e-16;
 Matches 1228: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTGCAGATTAATGCACTTCTCATACTGAGCGTACAGTATGAGAGGGGTATATC 60

Db 38201 ATCTGCAGATTAATGCACTTCTCATACTGAGCGTACAGTATGAGAGGGGTATATC 38260
 QY 61 ATGCGCGTGGATATTCGACGAGACCAAAAACAGAAATCCGGGTGGCCATACCCCG 120
 Db 38261 ATGCGCGTGGATATTCGACGAGACCAAAAACAGAAATCCGGGTGGCCATACCCCG 38320
 QY 121 GCGGCGTCCGGAACATACCCGTCGTCGATGAGTGTCTATCCAGGAGGTGGCGGA 180
 Db 38321 GCGGCGTCCGGAACATACCCGTCGTCGATGAGTGTCTATCCAGGAGGTGGCGGA 38380
 QY 181 GAGGCGTCCGGAACATACCCGTCGTCGATGAGTGTCTATCCAGGAGGTGGCGGA 240
 Db 38381 GAGGCGTCCGGAACATACCCGTCGTCGATGAGTGTCTATCCAGGAGGTGGCGGA 38440
 QY 241 GCGGCGTCCGGAACATACCCGTCGTCGATGAGTGTCTATCCAGGAGGTGGCGGA 300
 Db 38441 GCGGCGTCCGGAACATACCCGTCGTCGATGAGTGTCTATCCAGGAGGTGGCGGA 38500
 QY 301 GAATACGCGCGCTGCGACACGCGGAGATTTTCAGCGGCGGACGCACTGTGCGCAC 360
 Db 38501 GAATACGCGCGCTGCGACACGCGGAGATTTTCAGCGGCGGACGCACTGTGCGCAC 38560
 QY 361 CGTCTTGACACCGATGCGTGTGATTCGCGACACGTCATTTGCTACGAGACCGTC 420
 Db 38561 CGTCTTGACACCGATGCGTGTGATTCGCGACACGTCATTTGCTACGAGACCGTC 38620
 QY 421 CAGACCGCGGACGCGACATCCCTGTCGCGGATGAGCAAGTCGCGGTCGACATC 480
 Db 38621 CAGACCGCGGACGCGACATCCCTGTCGCGGATGAGCAAGTCGCGGTCGACATC 38680
 QY 481 GCGGCGCGGTTGGCGCTTACACCTGATCGCAACCAAGGGGCGCGGTGTGCTGATG 540
 Db 38681 GCGGCGCGGTTGGCGCTTACACCTGATCGCAACCAAGGGGCGCGGTGTGCTGATG 38740
 QY 541 GCGGCGCGGTTGGCGCTTACACCTGATCGCAACCAAGGGGCGCGGTGTGCTGATG 600
 Db 38741 GCGGCGCGGTTGGCGCTTACACCTGATCGCAACCAAGGGGCGCGGTGTGCTGATG 38800
 QY 601 TACAACGCGCGCGCATCGCCACAGCGGATGGGCGGACGCTTACGCTTACGATCAAC 660
 Db 38801 TACAACGCGCGCGCATCGCCACAGCGGATGGGCGGACGCTTACGCTTACGATCAAC 38860
 QY 661 ATGACCAACTTGGCGCACTGACGCGGAGTGTGCGGCGCGGATCCACATCCCTACTCA 720
 Db 38861 ATGACCAACTTGGCGCACTGACGCGGAGTGTGCGGCGCGGATCCACATCCCTACTCA 38920
 QY 721 TCGGCTACGAGCTCGAGGGTGGCGTCAACGTCGACCTGGTGGATGGGCGGCTCTG 780
 Db 38921 TCGGCTACGAGCTCGAGGGTGGCGTCAACGTCGACCTGGTGGATGGGCGGCTCTG 38980
 QY 781 GTGCCAGGCGCAAGGACCAACCAATTAAGTCTCGAATTCGTCGCGCATATGAACA 840
 Db 38981 GTGCCAGGCGCAAGGACCAACCAATTAAGTCTCGAATTCGTCGCGCATATGAACA 39040
 QY 841 GTGGCGTACGTGGATATATAGCCATGACAGGCGGCTTTCGAAGCTCACGACCG 900
 Db 39041 GTGGCGTACGTGGATATATAGCCATGACAGGCGGCTTTCGAAGCTCACGACCG 39100
 QY 901 ACCACCTACGACCAACCGGACGCTGCGGTCGACACACGCTGTTTACTCGTGGGCAAC 960
 Db 39101 ACCACCTACGACCAACCGGACGCTGCGGTCGACACACGCTGTTTACTCGTGGGCAAC 39160
 QY 961 ATGCGCGCTGCGTGGCGGAGAGAGCTGCACTACGCGCTGACCAACGCGAGATCCGTAT 1020
 Db 39161 ATGCGCGCTGCGTGGCGGAGAGAGCTGCACTACGCGCTGACCAACGCGAGATCCGTAT 39220
 QY 1021 GTGCTCGAGTTCGCGACACATGCGTGGGCGGCGGCGGCGGCGGATTCGCGACCTAGCC 1080
 Db 39221 GTGCTCGAGTTCGCGACACATGCGTGGGCGGCGGCGGCGGCGGATTCGCGACCTAGCC 39280
 QY 1081 AAAGGCTTTTGACGACGCAAGGGGCGTTACTGTCCGAACGCGGTGACGACGCTGGGG 1140

DB	39281	AAAGGCTTTTCACCGACGAGAAAGGGGGGCTTACTGTCCGAACGGGATGGCCACGACACTGGGG	39340
QY	1141	GTGCGGTTACCGAGCGCCGACGCGTCTGAGCTTGCAGCTCGCGGCTGTTACGCCAG	1200
Db	39341	GTGCGGTTACCGAGCGCCGACGCGTCTGAGCTTGCAGCTCGCGGCTGTTACGCCAG	39400
QY	1201	CACAGCTGGGAGTAAGGGAACGATGA	1228
Db	39401	CACAGCTGGGAGTAAGGGAACGATGA	39428

RESULT	6				
LOCUS	AB7613	1236 bp	DNA	PAT	22-JAN-2000
DEFINITION	Sequence 10 from Patent WO9836089.				
ACCESSION	AB7613				
VERSION	AB7613.1	GI:6736253			
KEYWORDS					
SOURCE	unidentified.				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 1236)				
AUTHORS	Flohe, U. and Singh, M.				
TITLE	TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE				
JOURNAL	Patent: WO 9836089-A 20-AUG-1998;				
FEATURES	FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)				
source	1. 1236				
	/organism="unidentified"				
	/db_xref="taxon:32644"				
BASE COUNT	236 a 395 c 385 g 220 t				
ORIGIN					

Query Match	99.9%	Score 1226.4;	DB 5;	Length 1236;
Best Local Similarity	99.9%	Pred. No. 7.2e-169;		
Matches 1277;	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0
QY	1	ATCTTCAGATTAAATCAATCTTCTTATACGTGAAAGCGTACAGATTCGAGAGGGTAATC	60	
Db	1	ATCTTCAGATTAAATCAATCTTCTTCTTCACTGAAAGCGTACAGATTCGAGAGGGTAATC	60	
QY	61	ATGCGGCTCGGTATTCGAGCCGAGACCAAAAACAAGAAATCCGGGTGGCATCACCCCG	120	
Db	61	ATGCGGCTCGGTATTCGAGCCGAGACCAAAAACAAGAAATCCGGGTGGCATCACCCCG	120	
QY	121	GCGGCGTCGCGGAAGTAACCCGTCGTGGCCATGAGGTGCTATCCAGGAGGTGGCGGA	180	
Db	121	GCGGCGTCGCGGAAGTAACCCGTCGTGGCCATGAGGTGCTATCCAGGAGGTGGCGGA	180	
QY	181	GAGGGCTCGGCTATCAACGACGCGATTTTCAGGCGGCGAGCGGCACTGTGCGCAC	240	
Db	181	GAGGGCTCGGCTATCAACGACGCGGATTTTCAGGCGGCGAGCGGCACTGTGCGCAC	240	
QY	241	GCGGACAGAGTGTGGGCGCGCGCTGATTTTCTCAAGGTCAAAAGCAAGCAATAGGGGG	300	
Db	241	GCGGACAGAGTGTGGGCGCGCGCTGATTTTCTCAAGGTCAAAAGCAAGCAATAGGGGG	300	
QY	301	GAATACGCGCCGCTGCGACACGGGCAAGATCTTGTACCTCTTGCAATTGAGCCCTGTCA	360	
Db	301	GAATACGCGCCGCTGCGACACGGGCAAGATCTTGTACCTCTTGCAATTGAGCCGCGTCA	360	
QY	361	CGAGCTGACACGATGAGTGTGATTCGGGACACGATTCGATTCGATTCGATTCGATTC	420	
Db	361	CGAGCTGACACGATGAGTGTGATTCGGGACACGATTCGATTCGATTCGATTCGATTC	420	
QY	421	CAGACGCGCGACGCGCACTACCCCTGCTTGGCCCGATGAGCAAGTGGCCGCTGACTC	480	
Db	421	CAGACGCGCGACGCGCACTACCCCTGCTTGGCCCGATGAGCAAGTGGCCGCTGACTC	480	
QY	481	GCGGCGGAGGTTGGGCTTACCACTGATGCGAAGCAAGGAGGCGCGGCTGTGCTATG	540	
Db	481	GCGGCGGAGGTTGGGCTTACCACTGATGCGAAGCAAGGAGGCGCGGCTGTGCTATG	540	

Oy	541	GGCGGGGTGCGCCGGCGTCGAAACCGCGCAACGTCGTGTGATGTGGCGCGCGACCGCGGGC	600
Db	541	GGCGGGGTGCGCCGGCGTCGAAACCGCGCAACGTCGTGTGATGTGGCGCGCGACCGCGGGC	600
Oy	601	TACAACGACGCGGCATCTGCAACGGGCATGTGGCGCGACCGTTACGGTCTTAGACATCAAC	660
Db	601	TACAACGACGCGGCATCTGCAACGGGCATGTGGCGCGACCGTTACGGTCTTAGACATCAAC	660
Oy	661	ATCGACAACCTTGGGCACTCTGACGCCGAGTCTGGGGCCG.ATCCACACTTCGCTACTCA	720
Db	661	ATCGACAACCTTGGGCACTCTGACGCCGAGTCTGGGGCCG.ATCCACACTTCGCTACTCA	720
Oy	721	TCGGCCTACGAGCTCGAGGGGTGCGCTTCGCAACGTCGCGACCTGGTATTGGGCGCGTCTG	780
Db	721	TCGGCCTACGAGCTCGAGGGGTGCGCTTCGCAACGTCGCGACCTGGTATTGGGCGCGTCTG	780
Oy	781	GTGCCAGGCGCCAAAGCACCACCAATTAGTCTGAATTCACTTGTGCGCATATGAACCA	840
Db	781	GTGCCAGGCGCCAAAGCACCACCAATTAGTCTGAATTCACTTGTGCGCATATGAACCA	840
Oy	841	GGTGGGACTAGTGGATATAGCCATTCGACACAGGGCGGCTTTGCAAGGCTCAGCAGCG	900
Db	841	GGTGGGACTAGTGGATATAGCCATTCGACACAGGGCGGCTTTGCAAGGCTCAGCAGCG	900
Oy	901	ACCACTACGACACCGCGAGCTTGCCCGTCGACGACACGAGCTTTTACTGCGTGGCGAC	960
Db	901	ACCACTACGACACCGCGAGCTTGCCCGTCGACGACACGAGCTTTTACTGCGTGGCGAC	960
Oy	961	ATGCGCGCGCTGGGCGGAGAGCGTGGACCTACGGCTTACCACAGCGCAGTGGCGGAT	1020
Db	961	ATGCGCGCGCTGGGCGGAGAGCGTGGACCTACGGCTTACCACAGCGCAGTGGCGGAT	1020
Oy	1021	GTGCTCGAGCTTGCCGACCATGCTGGCGGCGCGCTGGCGCTTCGATCTCGACTACCG	1080
Db	1021	GTGCTCGAGCTTGCCGCA.ATGCTGGCGGCGCGCTGGCGCTTCGATCTCGACTACCG	1080
Oy	1081	AAAGGCTCTGACGACGAAAGGGCGCTTACTCTCGAAGGGGTGGCCACGACGACTGGGG	1140
Db	1081	AAAGGCTCTTTCGACGACGAAAGGGCGCTTACTCTCGAAGGGGTGGCCACGACGACTGGGG	1140
Oy	1141	GTGCGCTTACCGGACCGCGCGAGCGTGTGGCTGTGACTCTGCGCGCTCTGTACCGCGAG	1200
Db	1141	GTGCGCTTACCGGACCGCGCGAGCGTGTGGCTGTGACTCTGCGCGCTCTGTACCGCGAG	1200
Oy	1201	CACACGTCGGGAGGTAAAGGAACGA.3A.1228	
Db	1201	CACACGTCGGGAGGTAAAGGAACGATTA.1228	
RESULT	7		
LOCUS	A89753	1236 bp	DNA
DEFINITION	Sequence 10 from Patent WO932862.		
ACCESSION	A89753		
VERSION	A89753.1	GI:6738287	
KEYWORDS			
SOURCE	unidentified.		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 1236)		
AUTHORS	Flohe,L. and Singh,M.		
TITLE	L-PLANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM		
JOURNAL	Patent: WO 9832862-A 30-JUL-1998;		
FEATURES	FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)		
source	Location/Qualifiers		
	1..1236		
	/organism="unidentified"		
	/db_xref="taxon:32644"		
BASE COUNT	236 a 395 c 385 g 220 t		
ORIGIN			

BASE COUNT	236 a	394 c	382 g	220 t	3 others
ORIGIN	/db_xref="taxon:33644"				
Query Match	99.7%	Score 1224.4	DB 5	Length 1235	
Best Local Similarity	99.8%	Pred. No. 1.4e-168			
Matches 1225, Conservative	0	Mismatches 3	Indels 0	Gaps 0	
QY	1	ATCTCGAGATTATCGAAGCTTTCTTCATCTCAAGAGTACACTATTCGAGAGGGGTAATC	60		
DB	1	ATCTTCGAGATTATTCGAACTTTCTTCACCTACACTAAGCTACAGATTCGAGAGGGGTAATC	60		
QY	61	ATGCGCGTGGTATTTCCGACCGAGACCMAAAACAGCAATTCGGGGTGCATACCCCG	120		
DB	61	ATGGCGCGTGGTATTTCCGACCGAGACCMAAAACAGCAATTCGGGGTGCATACCCCG	120		
QY	121	GCCGGCGTCCGGGAACTAACCCGTCGGGCAATGAGGTGCTCATCCAGGAGAGTGCAGGA	180		
DB	121	GCCGGCGTCCGGGAACTAACCCGTCGGGCAATGAGGTGCTCATCCAGGAGAGTGCAGGA	180		
QY	181	GAGGCGTGGCTATTCACCGACGCGAGATTTCAAGGCGGACGCGCAGCACTGTGTGGCACC	240		
DB	181	GAGGCGTGGCTATTCACCGACGCGAGATTTCAAGGCGGACGCGCAGCACTGTGTGGCACC	240		
QY	241	GCCGACGAGGTGTGGCCCGACGCTGATTTATTGCTCAAGGTCAAAGAACCGATAGGGGG	300		
DB	241	GCCGACGAGGTGTGGCCCGACGCTGATTTATTGCTCAAGGTCAAAGAACCGATAGGGGG	300		
QY	301	GAATACGGCGCGCTGCGGACAGGGCAGATCTTGTCAGCTTTGCAATTTGGCCGGCGTCA	360		
DB	301	GAATACGGCGCGCTGCGGACAGGGCAGATCTTTGTCAGCTTTGCAATTTGGCCGGCGTCA	360		
QY	361	CGTGGTTCACCGATGCGTTGTGGATTCCGGGACACAGCTCAATTGGCTACGAGACGCTG	420		
DB	361	CGTGGTTCACCGATGCGTTGTGGATTCCGGGACACAGCTCAATTGGCTACGAGACGCTG	420		
QY	421	CAGACCGCGGAGAGGGGCACTACCCGCTGTCCCGGATGAGGAGCGAAGTGGCGGATGACTC	480		
DB	421	CAGACCGCGGAGAGGGGCACTACCCGCTGTCCCGGATGAGGAGCGAAGTGGCGGATGACTC	480		
QY	481	GCGGCGCAAGTGTGGCGCTTACCACTGATGCGAACCMAAGGGGGCCGCGGTGTGCTGATG	540		
DB	481	GCGGCGCAAGTGTGGCGCTTACCACTGATGCGAACCMAAGGGGGCCGCGGTGTGCTGATG	540		
QY	541	GCGGGGGTGTCCCGGCTGTGAACCGGCGCAGCTGTGTGATCGGGCGCGGACCGCGCGG	600		
DB	541	GCGGGGGTGTCCCGGCTGTGAACCGGCGCAGCTGTGTGATCGGGCGCGGACCGCGCGG	600		
QY	601	TACAAAGCAGCGCGATCGGCAAGGAGCATGTGGGCGGAGACCGTTAGGTTCTAGACATCAAC	660		
DB	601	TACAAAGCAGCGCGATCGGCAAGGAGCATGTGGGCGGAGACCGTTAGGTTCTAGACATCAAC	660		
QY	661	ATGCAACAACTTCGGCAACTCGACCGCGAGTGTCTCGGGCCGGATCCACATCTGCTACTCA	720		
DB	661	ATGCAACAACTTCGGCAACTCGACCGCGAGTGTCTCGGGCCGGATCCACATCTGCTACTCA	720		
QY	721	TGGGCTTACGAGCTGAGAGGTGCGCTCAAGGTGCGGACGCTGGATTTGGGGCCGTCTG	780		
DB	721	TGGGCTTACGAGCTGAGAGGTGCGCTCAAGGTGCGGACGCTGGATTTGGGGCCGTCTG	780		
QY	781	GTCGACGAGCGCAAGGACCCAAATTAAGTCTCGAATTCACATTGTCGCGCATATGAAACA	840		
DB	781	GTCGACGAGCGCAAGGACCCAAATTAAGTCTCGAATTCACATTGTCGCGCATATGAAACA	840		
QY	841	GGTGGCGTACTGTTGATATTAACCATGACACAGGGCGGGGTGTTTGAGAGGCTCAGCACCG	900		
DB	841	GGTGGCGTACTGTTGATATTAACCATGACACAGGGCGGGGTGTTTGAGAGGCTCAGCACCG	900		
QY	901	ACCACTTAGACACACCGAGTTCGCGGTGACGACAGACGCTGTTTACTGCTGGGCGAAC	960		
DB	901	ACCACTTAGACACACCGAGTTCGCGGTGACGACAGACGCTGTTTACTGCTGGGCGAAC	960		

QY 961 ATGCCCCGCTGGTGGCCGAGAGAGCTGACCTAGCGGTGACCAACGACGATGCGGTAT 1020
 DB 961 ATGCCCCGCTGGTGGCCGAGAGAGCTGACCTAGCGGTGACCAACGACGATGCGGTAT 1020
 QY 1021 GTGCTGAGACTTGGCCGACCATGGCTGGCGGGGCGGTGGCGGTGCAATCCGGCTAGCC 1080
 DB 1021 GTGCTGAGACTTGGCCGACCATGGCTGGCGGGGCGGTGGCGGTGCAATCCGGCTAGCC 1080
 QY 1081 AAGAGCTTTGACAGCAGCAAGAGGGGCTTACTGCTCGAAGGGGTGGCCGACGCTGGGG 1140
 DB 1081 AAGAGCTTTGACAGCAGCAAGAGGGGCTTACTGCTCGAAGGGGTGGCCGACGCTGGGG 1140
 QY 1141 GTGCGCTTACCGAGAGCCCGCAGCGTGTGGCTGACTCTCGCGCGCTGTTACCGGAG 1200
 DB 1141 GTGCGCTTACCGAGAGCCCGCAGCGTGTGGCTGACTCTCGCGCGCTGTTACCGGAG 1200
 QY 1201 CACAGCTCGGAGTAAGGAGGATGA 1228
 DB 1201 CACAGCTCGGAGTAAGGAGGATGA 1228

RESULT 10
 LOCUS A87609 1235 bp DNA PAT 22-JAN-2000
 DEFINITION Sequence 6 from Patent WO9836089.
 ACCESSION A87609
 VERSION A87609.1 GI:6736249
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1235)
 FLOHE, L. and Singh, M.
 TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
 Patent: WO 9836089-A 20-Aug-1998;
 JOURNAL FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source Location/Qualifiers
 1..1235
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 235 a 395 c 384 g 220 t 1 others
 ORIGIN

Query Match 98.8%; Score 1213.4; DB 5; Length 1235;
 Best Local Similarity 99.8%; Pred. No. 5.4e-167;
 Matches 1225; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATCTGAGATTAATGAACTTCTTCATGAGCTAGACGTACAGTATCGAGGGGTAATC 60
 DB 1 ATCTGAGATTAATGAACTTCTTCATGAGCTAGACGTACAGTATCGAGGGGTAATC 60
 QY 61 ATGCGGCTGGATATCCGACGAGACCAAAACACAGAAATTCGGGTGGCCATCACCCG 120
 DB 61 ATGCGGCTGGATATCCGACGAGACCAAAACACAGAAATTCGGGTGGCCATCACCCG 120
 QY 121 GCGGGGTGCGGAACAACTAACCGTGGGCGCATGAGTGTCTATCCAGGAGGTGCCGA 180
 DB 121 GCGGGGTGCGGAACAACTAACCGTGGGCGCATGAGTGTCTATCCAGGAGGTGCCGA 180
 QY 181 GAGGGCTGGCTATACGACGAGGGAATTTAAGGCGGACGCGCAACTGTGGGACAC 240
 DB 181 GAGGGCTGGCTATACGACGAGGGAATTTAAGGCGGACGCGCAACTGTGGGACAC 240
 QY 241 GCGGACAGGTGTGGGCGGACGCTGATTTATGCTCAAGGTCAAAAGAACCGATAGCGCG 300
 DB 241 GCGGACAGGTGTGGGCGGACGCTGATTTATGCTCAAGGTCAAAAGAACCGATAGCGCG 300
 QY 301 GAATACGCGCGCTGAGACACGGGAGATTTGTCAGTTTGTGATTTGGCGGCTCA 360
 DB 301 GAATACGCGCGCTGAGACACGGGAGATTTGTCAGTTTGTGATTTGGCGGCTCA 360
 QY 361 CGTGCTTGAACGAGATGCTGTGGATTCGGGACCAACGATTAATGCTTACGAGACCGTC 420

DB 360 CGTGCTTGAACGAGATGCTGTGGATTCGGGACCAACGATTAATGCTTACGAGACCGTC 419
 QY 421 CAGACCGCGGACGCGCACTACCCCTGCTTCCCGGATGAGCGAAATGCGCGTCACTC 480
 DB 420 CAGACCGCGGACGCGCACTACCCCTGCTTCCCGGATGAGCGAAATGCGCGTCACTC 479
 QY 481 GCGGCGGAGTTGGGCTTACCATCTGATGCGAACCACAAAGGGGCGCGGTGTCTATG 540
 DB 480 GCGGCGGAGTTGGGCTTACCATCTGATGCGAACCACAAAGGGGCGCGGTGTCTATG 539
 QY 541 GCGGGGTGCCCCGCTGCAACCGGCGAGCTGTGTGATGCGCGCGGACCGCGCGC 600
 DB 540 GCGGGGTGCCCCGCTGCAACCGGCGAGCTGTGTGATGCGCGCGGACCGCGCGC 599
 QY 601 TACAAGCAGCGCGCATGCGCAAGGGCATGGGGGAGACGTTAGGTTCTAGACATCAAC 660
 DB 600 TACAAGCAGCGCGCATGCGCAAGGGCATGGGGGAGACGTTAGGTTCTAGACATCAAC 659
 QY 661 ATGCAAAACTTTCGCAACTGAGCGCGAGTTCGCGGCGGAGTCCACACTCGCTACTCA 720
 DB 660 ATGCAAAACTTTCGCAACTGAGCGCGAGTTCGCGGCGGAGTCCACACTCGCTACTCA 719
 QY 721 TCGGCTTACGAGCTGAGAGGTCCGTCAAACGTCGCCACCTGGGTATGGGGCGTCTG 780
 DB 720 TCGGCTTACGAGCTGAGAGGTCCGTCAAACGTCGCCACCTGGGTATGGGGCGTCTG 779
 QY 781 GTGCAGGCGGCAAGGCGCAACCAATTAATGCTGCAATTAATGCTGCGCATATGAACA 840
 DB 780 GTGCAGGCGGCAAGGCGCAACCAATTAATGCTGCAATTAATGCTGCGCATATGAACA 839
 QY 841 GGTGCGGTACTGTGGATATAGCATTCACCAAGGCGGCTGTTTGAAGGCTCACAGCG 900
 DB 840 GGTGCGGTACTGTGGATATAGCATTCACCAAGGCGGCTGTTTGAAGGCTCACAGCG 899
 QY 901 ACCACTTACGACACCGCGCGCTGCGGCGACGACGACGCTGTTTACGTTGGGAGCAG 960
 DB 900 ACCACTTACGACACCGCGCGCTGCGGCGACGACGACGCTGTTTACGTTGGGAGCAG 959
 QY 961 ATGCCCCGCTGGTGGCGGAAGAGCTGACCTAGCGGCTGACCAACGAGATGCGGTAT 1020
 DB 960 ATGCCCCGCTGGTGGCGGAAGAGCTGACCTAGCGGCTGACCAACGAGATGCGGTAT 1019
 QY 1021 GTGCTGAGCTTGGCCGACCATGGCTGGGGGCGGTGGCGGTGCAATCCGGCTAGCC 1080
 DB 1020 GTGCTGAGCTTGGCCGACCATGGCTGGGGGCGGTGGCGGTGCAATCCGGCTAGCC 1079
 QY 1081 AAGAGCTTTGACAGCAGCAAGGCGCTTACTGTGCGAAGGGGTGGCCAGCAGCTGGGG 1140
 DB 1080 AAGAGCTTTGACAGCAGCAAGGCGCTTACTGTGCGAAGGGGTGGCCAGCAGCTGGGG 1139
 QY 1141 GTGCGGTACCGGACCGCGGACGCTGCTGGCTTACTGCGGCGGTGTTAGCGCGAG 1200
 DB 1140 GTGCGGTACCGGACCGCGGACGCTGCTGGCTTACTGCGGCGGTGTTAGCGCGAG 1199
 QY 1201 CACAGCTCGGAGTAAGGAGGATGA 1228
 DB 1200 CACAGCTCGGAGTAAGGAGGATGA 1227

RESULT 11
 LOCUS A89749 1235 bp DNA PAT 22-JAN-2000
 DEFINITION Sequence 6 from Patent WO9832862.
 ACCESSION A89749
 VERSION A89749.1 GI:6738283
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1235)
 FLOHE, L. and Singh, M.
 TITLE L-7-ANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM

JOURNAL Patent: WO 9832862-A 30-JUL-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES Location/Qualifiers
 source 1.1235
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 235 a 395 c 384 g 220 t 1 others
 ORIGIN

Query Match 98.8%; Score 1213.4; DB 5; Length 1235;
 Best Local Similarity 99.8%; Pred. No. 5.4e-167;
 Matches 1225; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATCTTGAGATTAAATGCACTTTCTTCACTGTAAGCGTACAGTATGAGAGGGTATC 60
 DB 1 ATCTTGAGATTAAATGCACTTTCTTCACTGTAAGCGTACAGTATGAGAGGGTATC 60
 QY 61 ATGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 120
 DB 61 ATGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 120
 QY 61 ATGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 120
 DB 61 ATGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 120
 QY 121 GCCGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 180
 DB 121 GCCGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 180
 QY 121 GCCGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 180
 DB 121 GCCGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 180
 QY 181 GAGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 240
 DB 181 GAGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 240
 QY 181 GAGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 240
 DB 181 GAGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 240
 QY 241 GCCGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 300
 DB 241 GCCGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 300
 QY 301 GAATACGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 360
 DB 301 GAATACGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 360
 QY 301 GAATACGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 360
 DB 301 GAATACGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 360
 QY 361 CGTGTTCACGAGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 420
 DB 361 CGTGTTCACGAGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 420
 QY 361 CGTGTTCACGAGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 420
 DB 361 CGTGTTCACGAGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 420
 QY 421 CAGACGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 480
 DB 421 CAGACGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 480
 QY 421 CAGACGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 480
 DB 421 CAGACGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 480
 QY 481 GCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 540
 DB 481 GCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 540
 QY 481 GCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 540
 DB 481 GCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 540
 QY 541 GCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 600
 DB 541 GCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 600
 QY 541 GCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 600
 DB 541 GCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 600
 QY 601 TACAACGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 660
 DB 601 TACAACGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 660
 QY 601 TACAACGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 660
 DB 601 TACAACGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 660
 QY 661 ATGCAACAATTCGCGCACTGAGCGCGGATTCGCGCGGATTCGCGCGGATTCGCGCGG 720
 DB 661 ATGCAACAATTCGCGCACTGAGCGCGGATTCGCGCGGATTCGCGCGGATTCGCGCGG 720
 QY 661 ATGCAACAATTCGCGCACTGAGCGCGGATTCGCGCGGATTCGCGCGGATTCGCGCGG 720
 DB 661 ATGCAACAATTCGCGCACTGAGCGCGGATTCGCGCGGATTCGCGCGGATTCGCGCGG 720
 QY 721 TCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 780
 DB 721 TCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 780
 QY 721 TCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 780
 DB 721 TCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 780
 QY 781 GCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 840
 DB 781 GCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 840
 QY 781 GCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 840
 DB 781 GCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 840
 QY 841 GGTGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 900
 DB 841 GGTGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 900
 QY 841 GGTGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 900
 DB 841 GGTGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 900

QY 901 ACCACTAGACACCGGACGTTCCGCGTGCACAGACAGCCTGTTTACTGCGTGGCAAC 960
 DB 901 ACCACTAGACACCGGACGTTCCGCGTGCACAGACAGCCTGTTTACTGCGTGGCAAC 960
 QY 961 ATGCGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 1020
 DB 961 ATGCGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 1020
 QY 961 ATGCGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 1020
 DB 961 ATGCGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 1020
 QY 1021 GCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 1080
 DB 1021 GCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 1080
 QY 1021 GCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 1080
 DB 1021 GCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 1080
 QY 1081 GCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 1140
 DB 1081 GCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 1140
 QY 1081 GCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 1140
 DB 1081 GCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 1140
 QY 1141 GCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 1200
 DB 1141 GCGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 1200
 QY 1201 CACACGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 1260
 DB 1201 CACACGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 1260
 QY 1201 CACACGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 1260
 DB 1201 CACACGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 1260

RESULT 12

AB7611 1235 bp DNA PAT 22-JAN-2000
 LOCUS A87611
 DEFINITION Sequence 8 from Patent WO9836089.
 ACCESSION A87611
 VERSION A87611.1 GI:6736251
 KEYWORDS
 SOURCE
 ORGANISM
 unidentified.
 unclassified.

REFERENCE
 1 (bases 1 to 1235)
 Flohe, L. and Singh, M.
 TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
 Patent: WO 9836089-A 20-AUG-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)

FEATURES
 source
 Location/Qualifiers
 1.1235
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 235 a 394 c 385 g 220 t
 ORIGIN

Query Match 98.8%; Score 1212.8; DB 5; Length 1235;
 Best Local Similarity 99.8%; Pred. No. 5.6e-167;
 Matches 1225; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATCTTGAGATTAAATGCACTTTCTTCACTGTAAGCGTACAGTATGAGAGGGTATC 60
 DB 1 ATCTTGAGATTAAATGCACTTTCTTCACTGTAAGCGTACAGTATGAGAGGGTATC 60
 QY 61 ATGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 120
 DB 61 ATGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 120
 QY 61 ATGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 120
 DB 61 ATGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 120
 QY 121 GCCGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 180
 DB 121 GCCGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 180
 QY 121 GCCGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 180
 DB 121 GCCGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 180
 QY 181 GAGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 240
 DB 181 GAGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 240
 QY 181 GAGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 240
 DB 181 GAGGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 240
 QY 241 GCCGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 300
 DB 241 GCCGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 300
 QY 241 GCCGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 300
 DB 241 GCCGCGGTGGGTATTCGACGAGCAACCAAAACAAATCCGGGAGGCAATCCCGC 300

QY 301 GAATACGGCGCGCTGCGACAGGAGCATCTTCTCAGCTTCTTGCAATTTGGCCGCGTCA 360
 DB 301 GAATACGGCGCGCTGCGACAGGAGCATCTTCTCAGCTTCTTGCAATTTGGCCGCGTCA 359
 QY 361 CGTGTGACCGCATGCTGTTGGATTCGGCAGCACTGTAATTTCCCTACAGAGCCGTC 420
 DB 361 CGTGTGACCGCATGCTGTTGGATTCGGCAGCACTGTAATTTCCCTACAGAGCCGTC 419
 QY 421 CAAACGGCGAGGCGGACATACCCCTGCTTGGCCGATGAGCAAGTCCGCGTCACTC 480
 DB 420 CAAACGGCGAGGCGGACATACCCCTGCTTGGCCGATGAGCAAGTCCGCGTCACTC 479
 QY 481 GCGCGCCAGGTTGGCGCTTACACCTGATGCGAACCAGGGGGCCGCGTGTCTGATG 540
 DB 480 GCGCGCCAGGTTGGCGCTTACACCTGATGCGAACCAGGGGGCCGCGTGTCTGATG 539
 QY 541 GGGGGGTTGGCGCGCGTCAACCGGCGAGCTGTGTGATGCGGCGCGGACCGCGCG 600
 DB 540 GGGGGGTTGGCGCGCGTCAACCGGCGAGCTGTGTGATGCGGCGCGGACCGCGCG 599
 QY 601 TACAACGACCCGCGCATGCGCAACGGCATGGCGCGACCTTACAGTTCTAGACATCAAC 660
 DB 600 TACAACGACCCGCGCATGCGCAACGGCATGGCGCGACCTTACAGTTCTAGACATCAAC 659
 QY 661 ATGACAACTTGGGCACTGCAAGCGGAGTTCGCGCGCGGATCCACACTGCTACTCA 720
 DB 660 ATGACAACTTGGGCACTGCAAGCGGAGTTCGCGCGCGGATCCACACTGCTACTCA 719
 QY 721 TCGGCTACAGCTGCAAGGTTGCGTCAAAAGTGCAGCTGTGATTTGGGCGCTCTG 780
 DB 720 TCGGCTACAGCTGCAAGGTTGCGTCAAAAGTGCAGCTGTGATTTGGGCGCTCTG 779
 QY 781 GTGCGACGCGCGCAAGGCAACCAATTTAGTCTGCAATCACTTGCGCGCATATGAAGCC 840
 DB 780 GTGCGACGCGCGCAAGGCAACCAATTTAGTCTGCAATCACTTGCGCGCATATGAAGCC 839
 QY 841 GGTGCGTACTGCTGATATAGCCATGACGAGGCGGCTGTTTGAAGGCTCAAGACCG 900
 DB 840 GGTGCGTACTGCTGATATAGCCATGACGAGGCGGCTGTTTGAAGGCTCAAGACCG 899
 QY 901 ACCACCTACGACACCGGAGCTTGGCGTGCAGACACGCTGTTTACTGCTGGGCGAAC 960
 DB 900 ACCACCTACGACACCGGAGCTTGGCGTGCAGACACGCTGTTTACTGCTGGGCGAAC 959
 QY 961 ATGCGCGCTGCTGCGCAAGAGCTGACCTACGCGCTGACCAACGAGAGATCCGTTAT 1020
 DB 960 ATGCGCGCTGCTGCGCAAGAGCTGACCTACGCGCTGACCAACGAGAGATCCGTTAT 1019
 QY 1021 GTGCTGAGCTTGGCGACATGCTGGCGGCGGCGTGGCTGGAATCCGGCACTAGCC 1080
 DB 1020 GTGCTGAGCTTGGCGACATGCTGGCGGCGGCGTGGCTGGAATCCGGCACTAGCC 1079
 QY 1081 AAAGCTTTTGAAGGAGGAGGCGTTACTGTCGCAAGGAGGAGGAGGAGGAGGAGG 1140
 DB 1080 AAAGCTTTTGAAGGAGGAGGCGTTACTGTCGCAAGGAGGAGGAGGAGGAGGAGG 1139
 QY 1141 GTGCGCTTACGACGACCGCGAGCGTGTGCTGACTCTGCGCGCTGTTACGCGGAG 1200
 DB 1140 GTGCGCTTACGACGACCGCGAGCGTGTGCTGACTCTGCGCGCTGTTACGCGGAG 1199
 QY 1201 CACAGCTGGGAGTAAAGGAGGAGTGA 1228
 DB 1200 CACAGCTGGGAGTAAAGGAGGAGTGA 1227

RESULT 13

LOCUS A89751 1235 bp DNA PAT 22-JAN-2000
 DEFINITION Sequence 8 from Patent W09832862.
 ACCESSION A89751
 VERSION A89751.1 GI:6738285
 KEYWORDS
 SOURCE unidentified.

ORGANISM unidentified
 unclassified.
 REFERENCE 1 (bases 1 to 1235)
 AUTHORS Flohe, L. and Singh, M.
 TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
 JOURNAL Patent: WO 9832862-A 30-JUL-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source 1..1235
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 236 a 394 c 385 g 220 t
 ORIGIN

Query Match 98.8%; Score 1212.8; DB 5; Length 1235;
 Best Local Similarity 99.8%; Pred. No. 6.6e-167;
 Matches 1225; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATCTGCAAGATTATGCAACTTCTTCACTAGCAAGCGTATGAGAGGGGATATC 60
 DB 1 ATCTGCAAGATTATGCAACTTCTTCACTAGCAAGCGTATGAGAGGGGATATC 60
 QY 61 ATGCGCTGCGTATTCGACGAGACCAAAAAACAAGATTCCGGTGGCCATACCCCG 120
 DB 61 ATGCGCTGCGTATTCGACGAGACCAAAAAACAAGATTCCGGTGGCCATACCCCG 120
 QY 121 GCGGCGTGGCGCAACTAACCCGCTGTCGATGAGTGTCTATCCAGGCGAGTGGCGGA 180
 DB 121 GCGGCGTGGCGCAACTAACCCGCTGTCGATGAGTGTCTATCCAGGCGAGTGGCGGA 180
 QY 181 GAGGCGTGGCTATACCGACGCGGATTTCAAGGCGGCGGCGCAACTGTCGCGACC 240
 DB 181 GAGGCGTGGCTATACCGACGCGGATTTCAAGGCGGCGGCGCAACTGTCGCGACC 240
 QY 241 GCGGACAGGTGTGGCGCGACGCTATTTTGTCTCAAGGTCAAGAAACGATAGCGCG 300
 DB 241 GCGGACAGGTGTGGCGCGACGCTATTTTGTCTCAAGGTCAAGAAACGATAGCGCG 300
 QY 301 GAATACGGCGCGCTGCGACGAGGAGATCTTGTTCAGCTTCTTGCATTTGGCGCGTCA 360
 DB 301 GAATACGGCGCGCTGCGACGAGGAGATCTTGTTCAGCTTCTTGCATTTGGCGCGTCA 359
 QY 361 CGTGTGACCGCATGCTGTTGGATTCGGCAGCACTGTAATTTCCCTACAGAGCCGTC 420
 DB 361 CGTGTGACCGCATGCTGTTGGATTCGGCAGCACTGTAATTTCCCTACAGAGCCGTC 419
 QY 421 CAAACGGCGAGGCGGACATACCCCTGCTTGGCCGATGAGCAAGTCCGCGTCACTC 480
 DB 420 CAAACGGCGAGGCGGACATACCCCTGCTTGGCCGATGAGCAAGTCCGCGTCACTC 479
 QY 480 GCGCGCCAGGTTGGCGCTTACACCTGATGCGAACCAGGGGGCGGCTGTGATG 539
 DB 480 GCGCGCCAGGTTGGCGCTTACACCTGATGCGAACCAGGGGGCGGCTGTGATG 539
 QY 541 GCGGCGGTTGGCGCTTACACCTGATGCGAACCAGGGGGCGGCTGTGATG 600
 DB 541 GCGGCGGTTGGCGCTTACACCTGATGCGAACCAGGGGGCGGCTGTGATG 599
 QY 600 TACAACGACCCGCGCATGCGCAACGGCATGGGCGGACCGGTTTACACATCAAC 660
 DB 600 TACAACGACCCGCGCATGCGCAACGGCATGGGCGGACCGGTTTACACATCAAC 659
 QY 661 ATGACAACTTGGGCACTGCAAGCGGAGTTCGCGCGGATCCACACTGCTACTCA 720
 DB 660 ATGACAACTTGGGCACTGCAAGCGGAGTTCGCGCGGATCCACACTGCTACTCA 719
 QY 721 TCGGCTACAGCTGCAAGGTTGCGTCAAAAGTGCAGCTGTGATTTGGGCGCTCTG 780
 DB 720 TCGGCTACAGCTGCAAGGTTGCGTCAAAAGTGCAGCTGTGATTTGGGCGCTCTG 779
 QY 781 GTGCGACGCGCGCAAGGCAACCAATTTAGTCTGCAATTCATTTGCGCGCATATGAAGCA 840
 DB 781 GTGCGACGCGCGCAAGGCAACCAATTTAGTCTGCAATTCATTTGCGCGCATATGAAGCA 840

Db 780 GTGCCAGCGCCGCAAGGCAACCAATAGTGTGCAATTCACATTGTGCGCATATGAACA 839

QY 841 GGTGCGTACTGTGATATATACCATGACAGCGCGCTGTTTCGAAGCTCAGCACCG 900

Db 840 GGTGCGTACTGTGATATATAGCATGACAGCGCGCTGTTTCGAAGCTCAGCACCG 899

QY 901 ACCAGCTACGACGACCGGAGCTTGGCGTGACAGACGCTGTTTACTGCTGGGCAAC 960

Db 900 ACCAGCTACGACGACCGGAGCTTGGCGTGACAGACGCTGTTTACTGCTGGGCAAC 959

QY 961 ATGCGCGCTGCGTGGCGAAGACGTCACCTACGCGCTGACCAACGCGAGATGCGTAT 1020

Db 960 ATGCGCGCTGCGTGGCGAAGACGTCACCTACGCGCTGACCAACGCGAGATGCGTAT 1019

QY 1021 GTGCTGAGCTTGGCGACCAATGCTGGCGGCGCGCGCTGGTGAATCGGCACTAGCC 1080

Db 1020 GTGCTGAGCTTGGCGACCAATGCTGGCGGCGCGCGCTGGTGAATCGGCACTAGCC 1079

QY 1081 AAAGGTCTTTCGACGACGAGAGGCGCTTACTGTCCGAACGGGTGGCCACCGACCTGGG 1140

Db 1080 AAAGGTCTTTCGACGACGAGAGGCGCTTACTGTCCGAACGGGTGGCCACCGACCTGGG 1139

QY 1141 GTGCGCTTACCGAGCGCGCGCAAGCGTGTGCTGCTGACCTGCGCGCTCTTACCGGAG 1200

Db 1140 GTGCGCTTACCGAGCGCGCGCAAGCGTGTGCTGCTGACCTGCGCGCTCTTACCGGAG 1199

QY 1201 CACAGCTGGGAGTAAGGAGGAGCATGA 1228

Db 1200 CACAGCTGGGAGTAAGGAGGAGCATGA 1227

RESULT 14

LOCUS A87610 1229 bp DNA PAT 22-JAN-2000

DEFINITION Sequence 7 from Patent WO9836089.

ACCESSION A87610

VERSION A87610.1 GI:6736250

KEYWORDS

SOURCE unidentified.

ORGANISM unidentified.

REFERENCE 1 (bases 1 to 1229)

AUTHORS Flohe, L. and Singh, M.

TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE

JOURNAL Patent: WO 9836089-A 20-AUG-1998;

FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)

FEATURES

Location/Qualifiers

1..1229

Source

BASE COUNT 236 a 391 c 382 g 219 t 1 others

ORIGIN

Query Match 98.7%: Score 1211.8; DB: 5; Length 1229;

Best Local Similarity 99.7%: Pred. No. 9.3e-167; Mismatches 3; Indels 1; Gaps 1;

Matches 1224: Conservative 0;

QY 1 ATCTGCAATTAATGAACTTCTTCACTGAGAGCGTACAGTATCGAGAGGGTATC 60

Db 1 ATCTGCAATTAATGAACTTCTTCACTGAGAGCGTACAGTATCGAGAGGGTATC 60

QY 61 ATGCGGCTGATTCGAGCGAGCAAAACAAAGATTCGAGGTGAGCATCAACCCG 120

Db 61 ATGCGGCTGATTCGAGCGAGCAAAACAAAGATTCGAGGTGAGCATCAACCCG 120

QY 121 GCGCGGCTGCGGAACTAACCGCTGCGCATGAGTGTCTATCCAGCAGGTGCGGA 180

Db 121 GCGCGGCTGCGGAACTAACCGCTGCGCATGAGTGTCTATCCAGCAGGTGCGGA 180

QY 181 GAGGCTGCGCTATACAGCAGCGGATTCAGAGCGGCGGCAACTGGTGGGAC 240

Db 181 GAGGCTGCGCTATACAGCAGCGGATTCAGAGCGGCGGCAACTGGTGGGAC 240

QY 241 GCGGACAGGTGAGGCGGACGCTGATTTATTCGCAAGGTCAAGAACCGATACGGCG 300

Db 241 GCGGACAGGTGAGGCGGACGCTGATTTATTCGCAAGGTCAAGAACCGATACGGCG 300

QY 301 GAATACGCGCGCTGCGACACGAGGAAATTCCTTACGCTTCTGATTTGGCGCGTCA 360

Db 301 GAATACGCGCGCTGCGACACGAGGCG - GATCTTGTACGCTTCTGATTTGGCGCGTCA 359

QY 361 CGTCTTTCACAGCATGCTGTTGATTCGCGCAGCAGCTCAATTGGCTACAGACGCTC 420

Db 360 CGTCTTTCACAGCATGCTGTTGATTCGCGCAGCAGCTCAATTGGCTACAGACGCTC 419

QY 421 CAGACCGCGAGCGGACATACCGCTGCTTGGCGCGATGAGGAAGTGGCGGCTGACTC 480

Db 420 CAGACCGCGAGCGGACATACCGCTGCTTGGCGCGATGAGGAAGTGGCGGCTGACTC 479

QY 481 GCGCGCAGGTTGGCGCTTACCACTGATGCGAACCAGAGGCGCGGCTGCTGATG 540

Db 480 GCGCGCAGGTTGGCGCTTACCACTGATGCGAACCAGAGGCGCGGCTGCTGATG 539

QY 541 GCGCGGCTGCGCGGCTGCGAACCAGCGCGATGCTGATGCGGCGCGGCAACGCGGCG 600

Db 540 GCGCGGCTGCGCGGCTGCGAACCAGCGCGATGCTGATGCGGCGCGGCAACGCGGCG 599

QY 601 TACAGCGACCGCGCATGCGCAAGCGGATGCGGCGGCGGCTTACGCTTACACATCAAC 660

Db 600 TACAGCGACCGCGCATGCGCAAGCGGATGCGGCGGCGGCTTACGCTTACACATCAAC 659

QY 661 ATGACAACTTGGCGCATGCGACCGCGAGTTCTGGCGCGGATCCACCTGCTACTGA 720

Db 660 ATGACAACTTGGCGCATGCGACCGCGAGTTCTGGCGCGGATCCACCTGCTACTGA 719

QY 721 TCGGCTCAGAGCTCGAGGCTGCGCTCAACGCTGCGACCTGCTGATGAGGCGCGCTG 780

Db 720 TCGGCTCAGAGCTCGAGGCTGCGCTCAACGCTGCGACCTGCTGATGAGGCGCGCTG 779

QY 781 GTGCGAGCGCGCAAGGCTCAAAATAGTCTGCAATTCATTTGCGCGCATATGAACA 840

Db 780 GTGCGAGCGCGCAAGGCTCAAAATAGTCTGCAATTCATTTGCGCGCATATGAACA 839

QY 841 GGTGCGTACTGTGATATATAGCATGACAGAGGCGGCTGTTTCGAAGGCTACAGACG 900

Db 840 GGTGCGTACTGTGATATATAGCATGACAGAGGCGGCTGTTTCGAAGGCTACAGACG 899

QY 901 ACCAGCTACGACGACCGGAGCTTGGCGTGACAGACGCTGTTTACTGCTGGGCAAC 960

Db 900 ACCAGCTACGACGACCGGAGCTTGGCGTGACAGACGCTGTTTACTGCTGGGCAAC 959

QY 961 ATGCGCGCTGCGTGGCGAAGACGTCACCTACGCGCTGACCAACGCGAGATGCGTAT 1020

Db 960 ATGCGCGCTGCGTGGCGAAGACGTCACCTACGCGCTGACCAACGCGAGATGCGTAT 1019

QY 1021 GTGCTGAGCTTGGCGACCAATGCTGGCGGCGCGCGCTGGTGAATCGGCACTAGCC 1080

Db 1020 GTGCTGAGCTTGGCGACCAATGCTGGCGGCGCGCGCTGGTGAATCGGCACTAGCC 1079

QY 1141 GTGCGCTTACCGAGCGCGCGCAAGCGTGTGCTGCTGACCTGCGCGCTCTTACCGGAG 1200

Db 1140 GTGCGCTTACCGAGCGCGCGCAAGCGTGTGCTGCTGACCTGCGCGCTCTTACCGGAG 1199

QY 1201 CACAGCTGGGAGTAAGGAGGAGCATGA 1228

Db 1200 CACAGCTGGGAGTAAGGAGGAGCATGA 1227

RESULT 15

LOCUS A89750 1229 bp DNA PAT 22-JAN-2000

DEFINITION Sequence 7 from Patent WO9832862.
 ACCESSION A89750
 VERSION A89750.1 GI:6738284
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1229)
 FLOHE, L. and SINGH, M.
 TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
 JOURNAL Patent: WO 9832862-A 30-JUL-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source location/Qualifiers
 1..1229
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 236 a 391 c 382 g 219 t 1 others
 ORIGIN

Query Match 98.7%; Score 1211.8; DB 5; Length 1229;
 Best Local Similarity 99.7%; Pred. No. 9.3e-167;
 Matches 1224; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 ATCTTGCAATTAATCAACTTTCTTATCTAGAGC.ACAGTACGAGAGGGTAATC 60
 DB 1 ATCTTGAGATTAATCAACTTTCTTATCTAGAGCAGTACGAGAGGGTAATC 60
 QY 61 ATGCGCTCGTATTCGACGAGACCAAAACAAACGATTCGGGTGCGCATCACCCG 120
 DB 61 ATGCGCTCGTATTCGACGAGACCAAAACAAACGATTCGGGTGCGCATCACCCG 120
 QY 121 GCGGCGTCGCGGAGTAACCCGCTGCTGCGCATAGAGTCTCATCCAGGAGTCCGGA 180
 DB 121 GCGGCGTCGCGGAGTAACCCGCTGCTGCGCATAGAGTCTCATCCAGGAGTCCGGA 180
 QY 121 GCGGCGTCGCGGAGTAACCCGCTGCTGCGCATAGAGTCTCATCCAGGAGTCCGGA 180
 DB 121 GCGGCGTCGCGGAGTAACCCGCTGCTGCGCATAGAGTCTCATCCAGGAGTCCGGA 180
 QY 181 GAGGCGTCGCTATCACGAGCGGATTTCAAGCGGCGGAGCGGCAACTGTCGCGC 240
 DB 181 GAGGCGTCGCTATCACGAGCGGATTTCAAGCGGCGGAGCGGCAACTGTCGCGC 240
 QY 241 GCCGACGAGTGTGGGCGGAGCGTGAATTTATGCTCAAGAGTCAAAAGACGATAGCGG 300
 DB 241 GCCGACGAGTGTGGGCGGAGCGTGAATTTATGCTCAAGAGTCAAAAGACGATAGCGG 300
 QY 301 GAATACGCGCGCTTCGACAGCGGAGATCTGTCAAGTCTTCAAGTTCGCGCGCTCA 360
 DB 301 GAATACGCGCGCTTCGACAGCGGAGATCTGTCAAGTCTTCAAGTTCGCGCGCTCA 360
 QY 301 GAATACGCGCGCTTCGACAGCGGAGATCTGTCAAGTCTTCAAGTTCGCGCGCTCA 360
 DB 301 GAATACGCGCGCTTCGACAGCGGAGATCTGTCAAGTCTTCAAGTTCGCGCGCTCA 360
 QY 361 CGTCTTGACGATGCGTGTGATTCGCGACGAGTCAATTCCTAGAGAGCGGTC 420
 DB 361 CGTCTTGACGATGCGTGTGATTCGCGACGAGTCAATTCCTAGAGAGCGGTC 420
 QY 421 CAGACCGCGGAGCGGAGTACCCCTGCTTCCCGGATGAGCGAAGTCCGCGTCACTC 480
 DB 421 CAGACCGCGGAGCGGAGTACCCCTGCTTCCCGGATGAGCGAAGTCCGCGTCACTC 480
 QY 420 CAGACCGCGGAGCGGAGTACCCCTGCTTCCCGGATGAGCGAAGTCCGCGTCACTC 479
 DB 420 CAGACCGCGGAGCGGAGTACCCCTGCTTCCCGGATGAGCGAAGTCCGCGTCACTC 479
 QY 481 GCGGCGGAGTGTGGGCGTTCACACTGATGGAACCCAAAGGGGCGGCGTCTGATG 540
 DB 481 GCGGCGGAGTGTGGGCGTTCACACTGATGGAACCCAAAGGGGCGGCGTCTGATG 540
 QY 480 GCGGCGGAGTGTGGGCGTTCACACTGATGGAACCCAAAGGGGCGGCGTCTGATG 539
 DB 480 GCGGCGGAGTGTGGGCGTTCACACTGATGGAACCCAAAGGGGCGGCGTCTGATG 539
 QY 541 GCGGCGGAGTGTGGGCGTTCACACTGATGGAACCCAAAGGGGCGGCGTCTGATG 600
 DB 541 GCGGCGGAGTGTGGGCGTTCACACTGATGGAACCCAAAGGGGCGGCGTCTGATG 600
 QY 540 GCGGCGGAGTGTGGGCGTTCACACTGATGGAACCCAAAGGGGCGGCGTCTGATG 599
 DB 540 GCGGCGGAGTGTGGGCGTTCACACTGATGGAACCCAAAGGGGCGGCGTCTGATG 599
 QY 601 TACACGAGCGCGGATGCGCAAGCGCATGGCGGAGCGGATGAGTCTAGACATCAAC 660
 DB 601 TACACGAGCGCGGATGCGCAAGCGCATGGCGGAGCGGATGAGTCTAGACATCAAC 660
 QY 600 TACACGAGCGCGGATGCGCAAGCGCATGGCGGAGCGGATGAGTCTAGACATCAAC 659
 DB 600 TACACGAGCGCGGATGCGCAAGCGCATGGCGGAGCGGATGAGTCTAGACATCAAC 659
 QY 661 ATGCAACAATTCGCAACTGAGCGCGGATGCGGCGGATGCGCAACTGCGTACTCA 720
 DB 661 ATGCAACAATTCGCAACTGAGCGCGGATGCGGCGGATGCGCAACTGCGTACTCA 720
 QY 660 ATGCAACAATTCGCAACTGAGCGCGGATGCGGCGGATGCGCAACTGCGTACTCA 719
 DB 660 ATGCAACAATTCGCAACTGAGCGCGGATGCGGCGGATGCGCAACTGCGTACTCA 719
 QY 721 TCGGCTACGAGCTCGAGGCTGCGCTCAAAAGTCCGAGCTGCTGATGGGCGGCTCTG 780
 DB 721 TCGGCTACGAGCTCGAGGCTGCGCTCAAAAGTCCGAGCTGCTGATGGGCGGCTCTG 780

DB 720 TCGGCTACGAGCTCGAGGCTGCGCTCAAAAGTCCGAGCTGCTGATGGGCGGCTCTG 779
 QY 781 GTGCCAGGCGCCAGGCGCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 840
 DB 781 GTGCCAGGCGCCAGGCGCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 839
 QY 841 GGTGGGCTACTGTGTGATATAGCATGACAGGCGGCTGTTGAGAGGCTACAGACG 900
 DB 841 GGTGGGCTACTGTGTGATATAGCATGACAGGCGGCTGTTGAGAGGCTACAGACG 899
 QY 901 ACCACCTACGACACCCGACGCTTCCGCTGACAGCAGCAGCTGTTTACTGCGGAG 960
 DB 901 ACCACCTACGACACCCGACGCTTCCGCTGACAGCAGCAGCTGTTTACTGCGGAG 959
 QY 961 ATGCCGCGCTGCTGCGGAGAGCTGACCTAGCGGCTGACCAACGCGAGATGCGGTAT 1020
 DB 961 ATGCCGCGCTGCTGCGGAGAGCTGACCTAGCGGCTGACCAACGCGAGATGCGGTAT 1019
 QY 1021 GTGCTGAGCTTCCGACCATGCTGCGGCGGCGGCTGCGGCTGCAATCCGCACTAGCC 1080
 DB 1021 GTGCTGAGCTTCCGACCATGCTGCGGCGGCGGCTGCGGCTGCAATCCGCACTAGCC 1079
 QY 1081 AAAGGCTTTGACGCGCAGAGGCGGCTTACTGTCCGAAGGAGTGGCCAGCAGCTGGGG 1140
 DB 1081 AAAGGCTTTGACGCGCAGAGGCGGCTTACTGTCCGAAGGAGTGGCCAGCAGCTGGGG 1139
 QY 1141 GTGCGCTTACCGAGCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
 DB 1141 GTGCGCTTACCGAGCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1199
 QY 1201 CACAGCTGGGAGTAAGGAGGAGGATGA 1228
 DB 1201 CACAGCTGGGAGTAAGGAGGAGGATGA 1227

Search completed: June 22, 2000, 14:53:39
 Job time: 17634 sec


```

OY 235 GGCACCGCCGACAGGTGGGCGACGCTGATTATTGCTCAAGGTCAAAAGACGATA 294
    |||||||
DB 241 GGCACCGCCGACAGGTGGGCGACGCTGATTATTGCTCAAGGTCAAAAGACGATA 300
    |||||||
OY 295 GCGGCGGAATACGGCGCGCTGCGACACGGGCGAGATCTTGTACGTTCTTGTGATTTGGCC 354
    |||||||
DB 301 GCGGCGGAATACGGCGCGCTGCGACACGGGCGAGATCTTGTACGTTCTTGTGATTTGGCC 360
    |||||||
OY 355 GCGTACGCTGCTGACGCGATGCTGTTGGATTCGGGACGACGCTCAATTCCTTACGAG 414
    |||||||
DB 361 GCGTACGCTGCTGACGCGATGCTGTTGGATTCGGGACGACGCTCAATTCCTTACGAG 420
    |||||||
OY 415 ACCGTCCAGACCGCCGACGCGACCTACCCCTGCTTCCCGCATGACGCAAGTCCCGGT 474
    |||||||
DB 421 ACCGTCCAGACCGCCGACGCGACCTACCCCTGCTTCCCGCATGACGCAAGTCCCGGT 480
    |||||||
OY 475 CGACATCGCCGCGACGCTGCTTACCACTGATCGCAACCAAGGGGCGCGGTG 534
    |||||||
DB 481 CGACATCGCCGCGACGCTGCTTACCACTGATCGCAACCAAGGGGCGCGGTG 540
    |||||||
OY 535 CTGATGGGCGGGTGGCCGCGCTGCAACCGCGCGACGCTGCTGATCGCGCGCGAC 594
    |||||||
DB 541 CTGATGGGCGGGTGGCCGCGCTGCAACCGCGCGACGCTGCTGATCGCGCGCGAC 600
    |||||||
OY 595 GCGGCTATACGCAACCGCGCATCGCAACGGCGATGGCGCAACCGCTTACGCTTACGAC 654
    |||||||
DB 601 GCGGCTATACGCAACCGCGCATCGCAACGGCGATGGCGCAACCGCTTACGCTTACGAC 660
    |||||||
OY 655 ATCAACATGACAACTTGGGCACTGACGCGGAGTTCGGCGCGATTCACACTGCG 714
    |||||||
DB 661 ATCAACATGACAACTTGGGCACTGACGCGGAGTTCGGCGCGATTCACACTGCG 720
    |||||||
OY 715 TACTCATCGCCTACGACGCTCGAGGCTGCGTCAAACTGCGACGCTGATGGGCGC 774
    |||||||
DB 721 TACTCATCGCCTACGACGCTCGAGGCTGCGTCAAACTGCGACGCTGATGGGCGC 780
    |||||||
OY 775 GTCCTGCTGACAGGCGCCAAAGGCACTTATGCTCAATTCATTTGCGCGCATG 834
    |||||||
DB 781 GTCCTGCTGACAGGCGCCAAAGGCACTTATGCTCAATTCATTTGCGCGCATG 840
    |||||||
OY 835 AAACAGGTGCGGTGCTGATATAGCATGACGACGCGGCGCTGTTTCGAAGGCTCA 894
    |||||||
DB 841 AAACAGGTGCGGTGCTGATATAGCATGACGACGCGGCGCTGTTTCGAAGGCTCA 900
    |||||||
OY 895 CGACCGACACCTAGACACACCGGCTGCGCGTGCAGACGACGCTGTTTACTGCTG 954
    |||||||
DB 901 CGACCGACACCTAGACACACCGGCTGCGCGTGCAGACGACGCTGTTTACTGCTG 960
    |||||||
OY 955 GCGACATGCGCCGCTGCGGTGCGGAAGAGCTGACCTAGCGGCTGACCAACGCGAGATG 1014
    |||||||
DB 961 GCGACATGCGCCGCTGCGGTGCGGAAGAGCTGACCTAGCGGCTGACCAACGCGAGATG 1020
    |||||||
OY 1015 CCGTATGCTGCTGACGCTTGGCCACATGCTGCGGCGGCGCTGCGGCTGATTCGCGCA 1074
    |||||||
DB 1021 CCGTATGCTGCTGACGCTTGGCCACATGCTGCGGCGGCGCTGCGGCTGATTCGCGCA 1080
    |||||||
OY 1075 CTAGCCAAAGGCTTTCGACGACGCAAGGGGCGTTACTGTCGAAGGGGCGACGAC 1134
    |||||||
DB 1081 CTAGCCAAAGGCTTTCGACGACGCAAGGGGCGTTACTGTCGAAGGGGCGACGAC 1140
    |||||||
OY 1135 CTGCGGGTGGCGCTTACCGACGAGCGCGACGCTGCTGCTGCTGCTGCTGCTTAC 1194
    |||||||
DB 1141 CTGCGGGTGGCGCTTACCGACGAGCGCGACGCTGCTGCTGCTGCTGCTGCTTAC 1200
    |||||||
OY 1195 GCGGACGACGCTGCGGAGTAAAGGAAGGATGA 1228
    |||||||
DB 1201 GCGGACGACGCTGCGGAGTAAAGGAAGGATGA 1234
    |||||||

```

RESULT 2
 ID V49510 standard: DNA; 1260 BP.
 AC V49510;

```

DT 20-OCT-1998 (first entry)
DE Mycobacterium sp. AladH DNA.
KW Alanine dehydrogenase; AladH; ADH; diagnosis: tuberculosis; pathogen;
KW swimmers disease; vaccine; epidemic; infection; identification; ss.
OS Mycobacterium sp.
PN M09832862-A2.
PD 30-JUL-1998.
PE 29-JAN-1998; E00484.
PF 29-JAN-1997; EP-101339.
PA (FLOH/) FLOHE L. Kolk A. Singh M.
PI Fione L, Butler B, Kolk A. Singh M.
DR WPI: 98-427958/36.
PT Nucleic acid encoding alanine dehydrogenase of Mycobacterium marinum
PT - used for diagnosis of tuberculosis and other mycobacterial
PT diseases, also for treatment and prevention, for drug screening and
PT for bio-transformation
PS Disclosure, Page 11; 5pp; German.
CC This sequence encodes an alanine dehydrogenase (ADH) protein isolated
CC from a Mycobacterium sp. This protein is used to diagnose tuberculosis
CC and other mycobacterial infections (including 'swimmers' disease', caused
CC by M. marinum, a fish pathogen) in humans or animals. The protein can
CC also be used for control of epidemics and for vaccination, to screen for
CC agents with anti-mycobacterial activity, and in bio-transformations that
CC are specific for L-alanine. Also mycobacteria can be identified by
CC analysis of genomic ADH sequences. ADH is an antigen that is secreted
CC early during infection.
SQ Sequence 1260 BP; 243 A; 403 C; 389 G; 225 T;

Query Match 98.6%; Score 1210.4; DB 1; Length 1260;
Best Local Similarity 99.4%; Pred. No. 8e-244;
Matches 1227; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

OY 1 ATCTTCAGATTATGCACTTCTTATGATGAGCGTACAGTATGAGGAGTATC 60
    |||||||
DB 16 ATCTTCAGATTATGCACTTCTTATGATGAGCGTACAGTATGAGGAGTATC 75
    |||||||
OY 61 ATGGCGCTGCGATTTCGACCGACGACCAAAACAAGC-----AATTCGGGTGGCATC 114
    |||||||
DB 76 ATGGCGCTGCGATTTCGACCGACGACCAAAACAAGTTCATTCCTCGGATGCGCATC 135
    |||||||
OY 115 ACCCGCGCGCGCTGCGGAACTAACCGCTGCTGCGCATGAGTGTCTATCAGGACGAT 174
    |||||||
DB 136 ACCCGCGCGCGCTGCGGAACTAACCGCTGCTGCGCATGAGTGTCTATCAGGACGAT 195
    |||||||
OY 175 GCGGAGAGGCGCTGCGCTATACCGACGCGGATTTCAAGCGCGGCGCAACTGCTC 234
    |||||||
DB 196 GCGGAGAGGCGCTGCGCTATACCGACGCGGATTTCAAGCGCGGCGCAACTGCTC 255
    |||||||
OY 235 GGCACCGCCGACAGGTGGGCGACGCTGATTATGCTCAAGGTCAAAAGACGATA 294
    |||||||
DB 256 GGCACCGCCGACAGGTGGGCGACGCTGATTATGCTCAAGGTCAAAAGACGATA 315
    |||||||
OY 295 GCGGCGGAATACGGCGCGCTGCGACGCGGAGATCTTGTACGTTCTTGTGATTTGGCC 354
    |||||||
DB 316 GCGGCGGAATACGGCGCGCTGCGACGCGGAGATCTTGTACGTTCTTGTGATTTGGCC 375
    |||||||
OY 355 GCGTACGCTGCTGACGCGATGCTGTTGGATTCGGGACGACGCTCAATTCGCTTACGAG 414
    |||||||
DB 376 GCGTACGCTGCTGACGCGATGCTGTTGGATTCGGGACGACGCTCAATTCGCTTACGAG 435
    |||||||
OY 415 ACCGTCCAGACCGCCGACGCGACCTACCCCTGCTTCCCGCATGACGCAAGTCCCGGT 474
    |||||||
DB 436 ACCGTCCAGACCGCCGACGCGACCTACCCCTGCTTCCCGCATGACGCAAGTCCCGGT 495
    |||||||
OY 475 CGACTGCGCGCGCGAGTGGCGCTTACCACTGATGCAACCAAGGGGCGCGGTG 534
    |||||||
DB 496 CGACTGCGCGCGCGAGTGGCGCTTACCACTGATGCAACCAAGGGGCGCGGTG 555
    |||||||
OY 535 CTGATGGGCGGGTGGCGGCGTGAACGCGGCGAGCTGCTGATGATGCGCGCGGAC 594
    |||||||
DB 556 CTGATGGGCGGGTGGCGGCGTGAACGCGGCGAGCTGCTGATGATGCGCGCGGAC 615
    |||||||

```

QY 595 GCGGCTACACGACGACCCGATCGCAACGACGATGGCCGACCGTTACGGTTCTAGAC 654
 DB 616 GCGGCTACACGACGACCCGATCGCAACGACGATGGCCGACCGTTACGGTTCTAGAC 675
 QY 655 ATCAACATGACAACTTGGCACTGACGCGGATCTGGGGCCGATCCACACTGCC 714
 DB 676 ATCAACATGACAACTTGGCACTGACGCGGATCTGGGGCCGATCCACACTGCC 735
 QY 715 TACTCATCGGCTACGAGTTCGAGGTCGCTCAACGTCGCGACCTGGTATGGGGCC 774
 DB 736 TACTCATCGGCTACGAGTTCGAGGTCGCTCAACGTCGCGACCTGGTATGGGGCC 795
 QY 775 GTCTGTGTCGACGCGCCCAAGGACCCAAATTAGTCTGCAATTCTGCGGATATG 834
 DB 796 GTCTGTGTCGACGCGCCCAAGGACCCAAATTAGTCTGCAATTCTGCGGATATG 855
 QY 835 AAACGAGTGGCTACTGATGATATGATGATGATGATGATGATGATGATGATGATG 894
 DB 856 AAACGAGTGGCTACTGATGATGATGATGATGATGATGATGATGATGATGATG 915
 QY 895 CGACGAGTGGCTACTGATGATGATGATGATGATGATGATGATGATGATGATG 954
 DB 916 CGACGAGTGGCTACTGATGATGATGATGATGATGATGATGATGATGATGATG 975
 QY 955 GCGACATGCGCGCTGCGTGGTGGGAAAGAGTGTGATGATGATGATGATGATG 1014
 DB 976 GCGACATGCGCGCTGCGTGGTGGGAAAGAGTGTGATGATGATGATGATGATG 1035
 QY 1015 CCTATGCTGCTGAGTTCGCGACATGCTGCGGCGGCGGCGGCGGCGGCGGCGG 1074
 DB 1036 CCTATGCTGCTGAGTTCGCGACATGCTGCGGCGGCGGCGGCGGCGGCGGCGG 1095
 QY 1075 CTAGCCAAAGGCTTTCGACGACGAGGAGGCGTACTGTCCAGCGGTGGCCACGAC 1134
 DB 1096 CTAGCCAAAGGCTTTCGACGACGAGGAGGCGTACTGTCCAGCGGTGGCCACGAC 1155
 QY 1135 CTGGGGGTCGCTTACCGACGACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1194
 DB 1156 CTGGGGGTCGCTTACCGACGACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1215
 QY 1195 GCGGACGACGTCGCGGAGTAAGGAAGCATGA 1228
 DB 1216 GCGGACGACGTCGCGGAGTAAGGAAGCATGA 1249

RESULT 3
 ID V49625
 AC V49625:
 DT 20-NOV-1998 (first entry)
 DE Mycobacterium tuberculosis
 KW ss: Alanine dehydrogenase; tuberculosis; TB; mycobacteria; L-alanine.
 OS Mycobacterium tuberculosis.
 PN M09836089-A2.
 PD 20-JAN-1998.
 PR 28-JAN-1997; EP-101338.
 PA (FLOH/) FLOHE L.
 PI Flohe L, Hutter B, Kolk A, Singh M;
 DR WPI: 98-457123/39.
 PT Diagnosis of, e.g. tuberculosis from alanine dehydrogenase activity
 PT - useful for, e.g. for diagnosis, differentiation of strains,
 PT monitoring vaccination and identification of mycobacterial
 PT inhibitors
 PS Claim 13; Fig 2.3; 55pp; German.
 CC The Mycobacterium tuberculosis alanine dehydrogenase (AlaDH) is in
 CC the production of kits for diagnosing tuberculosis (TB) and other
 CC mycobacterial infections in humans or animals. Kits are used for direct
 CC diagnosis of TB on clinical samples (e.g. body fluids) and can
 CC differentiate between pathogenic and non-virulent strains, e.g. for
 CC identifying pathogens of the M. tuberculosis (M. t.) complex. The kit may
 CC also be used to identify substances that inhibit mycobacteria, for
 CC combatting epidemics and for vaccination follow-up. Oligonucleotides

CC derived from AlaDH are used similarly in diagnostic hybridisation tests,
 CC also for culture confirmation of isolated strains and for chromosome
 CC fingerprinting to detect/differentiate between mycobacteria, and for
 CC L-alanine-specific biotransformation reactions. Mycobacteria, and for
 CC antigen, present extracellularly after only a few days of growth, making
 CC it an ideal drug target.
 SO Sequence 1260 BP; 243 A; 403 C; 389 G; 225 T;

Query Match 98.6%; Score 1210.4; DB 1; Length 1260;
 Best Local Similarity 99.4%; Pred. No. 8e-244;
 Matches 1227; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 1 ATCTTGACGATTATGACACTTCTTCTACTGAGCTACAGTATGAGAGGGGTATTC 60
 DB 16 ATCTTGACGATTATGACACTTCTTCTACTGAGCTACAGTATGAGAGGGGTATTC 75
 QY 61 ATGCGCTGCGTATTCGACGACGACGACGACGACGACGACGACGACGACGACG 114
 DB 76 ATGCGCTGCGTATTCGACGACGACGACGACGACGACGACGACGACGACGACG 135
 QY 115 ACCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 174
 DB 136 ACCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 195
 QY 175 GCGGAGAGGCGCTGCGTATGACGACGACGACGACGACGACGACGACGACGACG 234
 DB 196 GCGGAGAGGCGCTGCGTATGACGACGACGACGACGACGACGACGACGACGACG 255
 QY 235 GCGGAGAGGCGCTGCGTATGACGACGACGACGACGACGACGACGACGACGACG 294
 DB 256 GCGGAGAGGCGCTGCGTATGACGACGACGACGACGACGACGACGACGACGACG 315
 QY 295 GCGGAGAGGCGCTGCGTATGACGACGACGACGACGACGACGACGACGACGACG 354
 DB 316 GCGGAGAGGCGCTGCGTATGACGACGACGACGACGACGACGACGACGACGACG 375
 QY 355 GCGTACGCTGCTTGCACGACGATCGTGTGATGATGATGATGATGATGATGATG 414
 DB 376 GCGTACGCTGCTTGCACGACGATCGTGTGATGATGATGATGATGATGATGATG 435
 QY 415 ACCGTGACAGCGCGGACGACGACGACGACGACGACGACGACGACGACGACGACG 474
 DB 436 ACCGTGACAGCGCGGACGACGACGACGACGACGACGACGACGACGACGACGACG 495
 QY 475 CGACTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 534
 DB 496 CGACTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 555
 QY 535 CTGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 594
 DB 556 CTGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 615
 QY 595 GCGGCTACACGACGACCCGATCGCAACGACGATGGCCGACCGTTACGGTTCTAGAC 654
 DB 616 GCGGCTACACGACGACCCGATCGCAACGACGATGGCCGACCGTTACGGTTCTAGAC 675
 QY 655 ATCAACATGACAACTTGGCACTGACGCGGATCTGGGGCCGATCCACACTGCC 714
 DB 676 ATCAACATGACAACTTGGCACTGACGCGGATCTGGGGCCGATCCACACTGCC 735
 QY 715 TACTCATCGGCTACGAGTTCGAGGTCGCTCAACGTCGCGACCTGGTATGGGGCC 774
 DB 736 TACTCATCGGCTACGAGTTCGAGGTCGCTCAACGTCGCGACCTGGTATGGGGCC 795
 QY 775 GTCTGTGTCGACGCGCCCAAGGACCCAAATTAGTCTGCAATTCTGCGGATATG 834
 DB 796 GTCTGTGTCGACGCGCCCAAGGACCCAAATTAGTCTGCAATTCTGCGGATATG 855
 QY 835 AAACGAGTGGCTACTGATGATGATGATGATGATGATGATGATGATGATGATG 894
 DB 856 AAACGAGTGGCTACTGATGATGATGATGATGATGATGATGATGATGATGATG 915

Qy	895	CGCCGACACACTTAAGCACCCGAGCTTGCGCTGACGAACGCTTTTACTGCGTG	954
Db	916	CGACCGACACACTTAAGCACCCGAGCTTGCGCTGACGAACGCTTTTACTGCGTG	975
Qy	955	GCGAACAATGCCCGCTGCTGCTGCGCCGAAGAGCTGCACCTACGCGCTACCAACGCAAGCATG	1014
Db	976	GCGAACAATGCCCGCTGCTGCTGCGCCGAAGAGCTGCACCTACGCGCTACCAACGCAAGCATG	1035
Qy	1015	CCGATATGCGCTCGAGCTCTTGCCGCACCAATGGCTGGCGCGCGCTGCGCGCTCGAATCCGCA	1074
Db	1036	CCGATATGCGCTCGAGCTCTTGCCGCACCAATGGCTGGCGCGCGCTGCGCGCTCGAATCCGCA	1095
Qy	1075	CTAAGCCAAAGTCTTTTGACAGCAGCAAGAGGGCGCTTACTTCGGAAGGGTGGCCACCGAC	1134
Db	1096	CTAAGCCAAAGTCTTTTGACAGCAGCAAGAGGGCGCTTACTTCGGAAGGGTGGCCACCGAC	1155
Qy	1135	CTGGGGGTGCCGTTACACCGAGCCCGCCAGCGTCTGGCTGACTCTCGGCGCGCTGTTAC	1194
Db	1156	CTGGGGGTGCCGTTACACCGAGCCCGCCAGCGTCTGGCTGACTCTCGGCGCGCTGTTAC	1215
Qy	1195	GCCGAGCAGACGTGCGGGAGTAAAGGAAGCATGA	1228
Db	1216	GCCGAGCAGACGTGCGGGAGTAAAGGAAGCATGA	1249

	RESULT	4	
VQ9511	ID	VQ9511 standard; DNA; 682 BP.	
AC	VQ9511	standard; DNA; 682 BP.	
DT	20-OCT-1998	(first entry)	
DE	Mycobacterium marinum Mar3	DNA.	
KW	Alanine dehydrogenase; AlADH; ADH; diagnosis; tuberculosis; pathogen;		
OS	swimmers disease; vaccine; epidemic; infection; identification; ss.		
FH	Mycobacterium marinum.		
FT	Key	Location/Qualifiers	
FT	CDS	1..682	
FT		/tag= a	
FT		/codon_start= 3	
FT		/product= "Mar3"	
FT		/note= "Alanine dehydrogenase"	
PN	MO9832862-A2.		
PD	30-JUL-1998		
PR	29-JAN-1998; E00484		
PF	29-JAN-1997; EP-101339.		
PA	(FLOH/) FLOHE L.		
PI	Flohe L, Hutter B, Kolk A, Singh M;		
PJ	WPI: 98-427958/36.		
DR	P-PGDB: W64481.		
PT	Nucleic acid encoding alanine dehydrogenase of Mycobacterium marinum		
PT	- used for diagnosis of tuberculosis and other mycobacterial		
PT	diseases, also for treatment and prevention, for drug screening and		
PS	for bio-transformation		
PS	Claim 1: Page 34-35; 57pp; German.		
CC	This sequence encodes an alanine dehydrogenase (ADH) protein, Mar3		
CC	isolated from Mycobacterium marinum. This protein is used to diagnose		
CC	tuberculosis and other mycobacterial infections (including 'swimmers'		
CC	disease, caused by M. marinum, a fish pathogen) in humans or animals		
CC	The protein can also be used for control of epidemics and for		
CC	vaccination, to screen for agents with anti-mycobacterial activity, and		
CC	in bio-transformations that are specific for L-alanine. Also mycobacte-		
CC	ria can be identified by analysis of genomic ADH sequences. ADH is an anti-		
CC	that is secreted early during infection.		
SQ	Sequence 682 BP; 105 A; 254 C; 225 G; 98 T;		

Query Match	38.1%	Score 467.6	DB 1	Length 682
Best local Similarity	80.4%	Pred. 2.9e-89		
Matches 568	Conservative	0	Mismatches 134	Indels 0
Gaps				
QY	98	AATTCGGGTGGCATCA	●	CCCCGGCGGCGCGGAACTAACCCGTCGCGCATGAGG 157
DB	1	AATTCGGGTAGGATACACCCCGCGGGGTGCGCCCTTACCAACGCGGCGCCAGAG 60		

QY	158	TGCTCATCAGGCGAGGTGGCGAGAGGGCTGCGCTCTACCGACGGGGATTCTAAAGCGG	21
Db	61	TGCTGATCTCAGAGCCGGTGGCCGAGAAAGGCTCCGCCCATCTCCGACGGCGAGCTTAAGGCGG	120
QY	218	CAGGCGCGCACTGTTGGCGCACCGCCGACAGGTGGGGCCGCGGTGATTATCTCTCA	277
Db	121	CCGGTGGCCACTATATACACCGCCGACCAAGGTGGTGGGGATGCGGACCTCTCTCTCA	180
QY	278	AGGTCAAAAGAACCGATAGCGCGGGAATACGGCCGCTGCGACACAGGGAGAGATTCTTTCA	337
Db	181	AGGTCAAAAGAACCCATGATGAGTCCGAGTACGGCCGGGCTGCCCGGGGCGACACCTTTTCA	240
QY	338	CGTCTTTCATTTTGGCGCGTCAAGTGTCTTGAACCGATGGTGTGGATTCCGGGACCA	397
Db	241	CTTACTCTCACTTGGCGGCTCTCGCGCCCTTGCACCGATGCTCTGTGAAGTCCGGGACCA	300
QY	388	CGTCAATTGCCTACGAGT CGTCCAGACCGCGCGACGCGCACTTACCCCTGCTTGGCCCCA	457
Db	301	CGTCCATGCGCTACGAGACGGGTGCAGACCGCGACCGCGATTCGCGCTCTG6CCCCCA	360
QY	458	TGAGGGAAGTGGCGGGTGAATCGCCGCCCGCAGGTGCGGTACGACCTGATCGAAGCC	517
Db	361	TGAGGGAAGTGGCGGGCGCGGCGCTGTCCGCCCAAGTGGGGCTTACCACTATATGGGACCG	420
QY	518	AAGGGGCGCGGGTGTGCTGATGAGGGCGGGTGTCCCGCGCTGAACCGCGCGAGTGTGTG	577
Db	421	ACGGGCGTTCGGCGCGCTGCTGATGGGCGCGGTGCCCGCGTCAACCCGTCCGACGTGTGTG	480
QY	578	TGATGAGGCGCGGCGACCGCCGGGTACAAAGCGAGGCCCATGCGCAACAGCGGATGGGCGCA	637
Db	481	TGATGAGGCGCGGCGACCGCGGGATTAACAAAGCGCGCGCGGTGGGCCCAAGGCGATGGGCGCA	540
QY	638	CCGTTACGGTTCATGACATCAACATCGACAACATTGGGCAACTGGAGCCGAGTCTCGG	697
Db	541	TGGTACCGGTGTGGATGTCAACATCAACAGCTCCGCGCAATGACGCGGAGTGTGGCG	600
QY	698	GCGGATCCACACTGCTACTATCATGAGCTTACGAGCTGAGAGGTGCGCGTCAAAAGTGGCG	757
Db	601	GTCGGGTCCGAGACCGCGTACTGCTGAGCCCTGCACTTGAGAGATGCGGACATCCAGCGCG	660
QY	758	ACCTGTGATTGGGCGCTGCT 779	
Db	661	ACATGTGATCGGGGCGCTCT 882	

RESULT 5
N91423
ID N91423; standard; DNA; 1125 BP.
AC N91423;
DT 01-FEB-1991 (first entry)
DE Sequence of heat-resistant alanine dehydrogenase (AH) gene with mol. wt.
DE 2 MD or less
KW Enzyme; ds.
OS *Bacillus stearothermophilus* IFO 12550.
PN T01043194-A.
PD 15-FEB-1989.
PF 10-AUG-1987; 200524.
PR 10-AUG-1987; JP-200524.
PA (NIRA) Unilika KK.
DR WPI: 89-096096/13.
PT Recombinant plasmid for transforming *Escherichia coli* -
PT obtd. by connecting heat resistant alanine dehydrogenase gene to
PT vector plasmid
PS Disclosure; Fig 2, p693; 9pp; Japanese.
CC A recombinant plasmid contg. heat-resistant alanine dehydrogenase (AH)
CC gene with mol. wt. 2 or less MD is claimed. Cells transformed with the
CC vector produce high levels of heat-resistant AH. Transformed *E. coli*
CC is useful as a clinical inspection reagent.
SQ Sequence 1125 BP; 246 A; 295 C; 367 G; 217 T;

Query Match	22.68;	Score 277.2;	DB 1;	Length 1125;
Best Local Similarity	56.28;	Pred. NO. 1.3e-49;		

Matches 604; Conservative 0; Mismatches 458; Indels 12; Gaps 4;

```

61 ATGCGGCTGGTATTCGCGAGACCAACAAACGAAATTCGGGTGGCCATCACC 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1 ATGAGATGCGATTCGCAAAAGAAATCAAAACAAATGAAACCGCTCCGCTAC 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 GCGGCGTGGCGAACTAACCCGTCGTCGCAATGAGTGTCT--CATCCAGCAGT 177
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61 GCGAGGCTGATGCGTGTCAAAAGGGGGGCAATGAGTGTGATGTGAGACGAA 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
178 GGAAGAGGCTGCGCTATTCACCGACGCGGATTTCAAGCGGCGAGCGCCGCA 237
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 GCGGCTGCTGCGGGGTTTCCGATTCGAGTAAAGAAACCGGCGAGCTGATC 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
238 ACCGCGCAGAGTGTGGCGCGCGCTGATTTATGCTCAAGTCAAGACGATAG 297
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
181 CGAGCTGGAGAGATCTTGGAGCGGGGAGATGCTGTGAAGTGAAGAGCGCG 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
298 GCGGAATACGCGCGCTGCGACACGCGGAGATCTTGTACGTTCTTGATTTG 357
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
241 CGAGATTCGCTATTTTGGCCCGGATTTGTTTACGATTTTGTATTTAGCG 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
358 TCACGCTGTGACGAGTGTGTTGATTCGGACCA .GTCATTTGCTACGAGAC 117
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
301 GCGGAAGCGCTACGAAAGCGCTGTCGACAAAAGTGTGCGCATCGTTACGAG 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
418 GTCGACAGCGCGCGCGCCTGCGACACGCGGAGATCTTGTGAGCGAGTGC 477
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
361 GTCGACTGCGGAGAGCGCTGCGACCTGTGACCGCGATGAGTACGAGTGC 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
478 CTGCGCGCCAGTGTGCGCTTACACCTGATCGAAACCAAGGGGCGCGGTG 537
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
421 ATCTCGTGTGAACTGCGCGCCAGTTTCTGAAAGCGCGGAGGAAAGGCA 480
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
538 ATGCGGGGGTGTGCGCGCGCTGCAACCGCGAGTGTGATGCGCGCGCACG 597
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
481 CTGCGCGCGCTGCGCGGAGTGTGCGCGCGGCGGAAAGTACG/TACTCG 540
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
598 GCGTACACGACGCGCGCGCTGCGCAACGCGATGCGCGCGCTTACGATC 657
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
541 GCGAGCAACGCGCGGAAATCGGCTGCGTCTGCGTACGCGTACGATTTG 600
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
658 AACATGCAAACTTGGCGCACTGCGAGCGAGTGTGCGCGGAGTCCACAT 717
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
601 AACGCGAGCGCTGCGCGAGCTGATGATTTGTCGCGACACGTCGACGAG 660
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
718 TCATCGGCTACGAGCTGAGAGGTGCGTCAAACTGCGCTGATTTGAGCG 777
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
661 TCAGACTGTCATACATATGCGCGAGTGTGCGGAAATCGATTTGGTCTG 720
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
778 CTGCTGCGAGCGCGCAAGCAACCAATTTAGTCTCGAATTCATCTTCTG 837
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
721 TTGATCCCGGGGGGGA--AGCGAAGCTGTGTGACGGAAGATGGTCTG 777
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
838 CAGGCTGCTAGCTGATGATATGATGATGATGATGATGATGATGATGAT 894
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
778 CCGGATCGGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 837
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
895 CGAGCGACCACTACGACACCGCGAGTGTGCGCGTACGACACGCTTTT 954
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
838 CGAGCGTACGACGAGTGTGATGATGATGATGATGATGATGATGATG 897
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
955 GCGAAGTGTGCGGCTGCGTGTGCGGAAAGTGTGATGATGATGATGATG 1014
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
898 GCGAAGTGTGCGGCTGCGTGTGCGGAAAGTGTGATGATGATGATGATG 954
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1015 CCGATGTGCTGAGCTGCGTGTGACATGCTGCGGCGGCGGCTGCGTGA 1074
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
955 CCGATGTGCTGAGCTGCGTGTGACATGCTGCGGCGGCGGCTGCGTGA 1014
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1075 CTGACCAAGGCTTTTGGAGCGGAGAGGCGGCTTACTGTCGGAAGGCG 1128
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1015 CTGTTAAAGGATCAACACGCTGACGCGGCAATCGTACGAAAGGCTG 1068
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 6
V52155/C
ID V52155 standard; DNA: 28171 BP.

AC V52155;
DE 23-OCT-1998 (first entry)
DE Streptococcus pneumoniae genome fragment SEQ ID NO:22.
KW Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
KW computer readable medium; vaccine; pharmaceutical composition; ds.
OS Streptococcus pneumoniae.
PN M09618931-A2.

PD 07-MAY-1998;
PF 30-OCT-1997; U19588.
PR 31-OCT-1996; US-029960.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M,
PI Kunsch CA, Rosen CA;
PI WPI: 98-27225/24.

PT Computer-readable medium with recorded Streptococcus pneumoniae
PT polynucleotide sequences - useful in diagnostic kits and assays, and
PT pharmaceutical compositions and vaccines for Streptococcus
PT pneumoniae

PS Claim 1: Page 273-289; 1409pp; English.

CC The present invention describes a computer readable medium which has
CC the nucleotide sequences SEQ ID NO:1 to 391 (V52134 to V52524) recorded
CC on it, or a representative fragment or a sequence at least 95% identical
CC to SEQ ID NO:1 to 391. The nucleotide sequences depicted in SEQ ID NO:1
CC to 391 (V52134 to V52524) are genomic fragments from Streptococcus
CC pneumoniae. The present invention also describes an isolated nucleic acid
CC molecule encoding a homologue of any of the fragments of the S.pneumoniae
CC genome (SEQ ID NO:1 to 391) where the nucleic acid molecule is produced
CC by a process comprising: (a) screening a genomic DNA library using as a
CC probe a target sequence defined by any of the sequences in SEQ ID NO:1
CC to 391, identifying members of the library which contain sequences
CC that hybridize to the target sequence and isolating the nucleic acid
CC molecules from the members; or (b) isolating mRNA, DNA or cDNA produced
CC from an organism, amplifying nucleic acid molecules whose nucleotide
CC sequence is homologous to amplification primers derived from the
CC fragment of the S. pneumoniae genome to prime the amplification and
CC isolating the amplified sequences. The computer readable medium can be
CC used in a computer-based system for identifying fragments of the
CC S. pneumoniae genome of commercial importance, or expression modulating
CC fragments of the S. pneumoniae genome. Products from the present
CC invention can be used in diagnosis kits and assays, and pharmaceutical
CC compositions and vaccines for S. pneumoniae.
SQ Sequence 28171 BP; 8507 A; 5080 C; 6440 G; 8144 T;

Query Match 17.3%; Score 212.4; DB 1; Length 28171;

Best Local Similarity 53.9%; Pred. No. 5.1e-36;
Matches 533; Conservative 0; Mismatches 436; Indels 19; Gaps 4;

```

61 ATGCGGCTGGTATTCGCGAGACCAACAAACGAAATTCGGGTGGCCATCACC 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
21711 ATGTAATGGAATCCCAAGAAATTAATAAGAAACGATGCTGCCCTACACCT 21652
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 GCGGCGTGGCGGAACTAACCCGTCGTCGCAATGAGTGTCT--CATCCAGCAGT 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
21651 GCAGGCTGTCTATGCTTACTGATGCTGTCATGCTGCTTATCAAAATGCTGT 21592
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
181 GAGGCTGCGCTATTCACCGACGCGGATTTCAAGCGGCGCGCACTGTGCGCAC 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
21591 CTGCGTTTGTGCTTACTGATGCTGCTACTCAAAAGAGAGCTGATTTGCTGCTACT 21532
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
241 GCGGACAGAGTGTGGGCGGACGCTGATTTATGCTCAAGGTCAAAAGACGATAGG 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
21531 GCTGTGAAACCTTGGGC--AGCAGAGTGTGTGTGAAAGTAAAGAACTTAACTTCT 21475
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
301 GAATAGCGCGCGCTGCGACACGCGGAGATGCTTGTGACGTTCTTGGATTGGCGCGTCA 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
21474 GAATAGCTTACTTGGCGGACGATCTTCTTCTTCACTACTTGCACATGCGCGTGTCT 21415
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

QY 361 CGTGGTGGACCGATGCGTGTGGATTCGGACACGTAATGCTCCTAGAGACGCTC 420
 DB 21414 CAGAGATTAGCAGAGATGATGTT-----AACGCAAAAAAACAAGTAATGTT 21367
 QY 421 CAGACCGCGGACGACGATACCCCTGCTGGCCGATGAGGAGGAGTGGCGGTGACTC 480
 DB 21366 CGTGACATTCAGAGACAACTACCGGCTCGTCTATGAGAGTGGAGTGGAGTGGATG 21307
 QY 481 GCCGCCGAGTGGGCGTTACCACTGATGCAACCAAGGAGGCGCGGTGCTGATG 540
 DB 21306 GCTGTTCAAATGAGAGCTACTCTCTACTAAGCAAGTGTGGCTGTGTTCTACTT 21247
 QY 541 GGGGGGGTGGCGGCTGCAACGCGCGGAGTGTGGTGGGCGCGCGCGCGCGCG 600
 DB 21246 GGTGTGTACAGAGTGTTCACAAAGAGAAAGTAATCATCATGCGTGGTGTGCGGT 21187
 QY 601 TACACGACGCGCGGATGCGCAACGCGATGGCGCGCGCGTACGTTACATCAAC 660
 DB 21186 ACACATGCTGCGCGGATGCGCGCGGATGCGCGGATGCGCGGATGCGCGGAT 21127
 QY 661 ATCGCAAAATTCGCGGATGCGCGGATGCGCGGATGCGCGGATGCGCGGATGCG 720
 DB 21126 TCCACGCTGCTCTCACTTCTAGAGAGATCTTGGAGATC AATTCAACTCTTATGCT 21067
 QY 721 TCGGCTACGAGCTGAGGAGTGGCGGATGCGCGGATGCGCGGATGCGCGGATGCG 780
 DB 21066 AATTCATTCAAATGAGCAAGATGAGAGATGAGATGAGATGAGATGAGATGAG 21007
 QY 781 GTGCGGAGGCGCGGATGCGCGGATGCGCGGATGCGCGGATGCGCGGATGCGCGG 840
 DB 21006 ATCCCTGCTGCGGATGCGCGGATGCGCGGATGCGCGGATGCGCGGATGCGCGG 20947
 QY 841 GTGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGG 897
 DB 20946 GGTGTGTACAGAGTGTTCACAAAGAGTGGAGTGGAGTGGAGTGGAGTGGAG 20888
 QY 898 CCGACACCTACGACCAACCGGATGCGCGGATGCGCGGATGCGCGGATGCGCGG 957
 DB 20887 GTACACAGGACGATGACGCGGATGCGCGGATGCGCGGATGCGCGGATGCGCGG 20828
 QY 958 AACATGCGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGG 1017
 DB 20827 AATATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 20768
 QY 1018 TATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1045
 DB 20767 TATATGAAAGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 20740
 RESULT 7
 V74442/C
 ID V74442 standard; DNA: 9280 BP.
 AC V74442;
 DT 16-MAR-1999 (first entry)
 DE Staphylococcus aureus contig SEQ ID #131.
 KW Computer readable medium; vaccine; S. aureus infection; immunodetection;
 KW Cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KW Skin infection; surgical wound infection; scalded skin syndrome;
 KW Toxic shock syndrome; ds.
 OS Staphylococcus aureus.
 FH Key Location/Qualifiers
 FT FT 841..900
 FT FT misc_feature
 FT FT /tag-a
 FT FT /note-
 FT FT "these bases represent a line of missing text in
 FT FT the sequence listing in the specification. They
 FT FT are included to maintain the nucleotide numbering
 FT FT given in the specification for this DNA sequence"
 FT FT misc_feature 2641..2700
 FT FT /tag-b
 FT FT /note-
 FT FT "these bases represent a line of missing text in
 FT FT the sequence listing in the specification. They
 FT FT are included to maintain the nucleotide numbering
 FT FT given in the specification for this DNA sequence"

FT misc_feature 4441..4500
 FT /tag-c
 FT /note-
 FT "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"
 FT misc_feature 6241..6300
 FT /tag-d
 FT /note-
 FT "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"
 FT misc_feature 8041..8100
 FT /tag-e
 FT /note-
 FT "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"
 EP-786519-A2.
 PD 30-JUL-1997.
 PF 07-JAN-1997: 100117.
 PR 05-JAN-1996: US-009861.
 PA (HMAN-) HUMAN GENOME SCI INC.
 PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA,
 PI Rosen CA;
 DR WPI: 97-374922/35.
 PT Polynucleotide(s) and proteins derived from Staphylococcus aureus
 PT stored on computer readable medium and used in the production of
 PT anti-S. aureus vaccines
 PS Claim 1: Page 710-715: 3271pp. English.
 CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
 CC of the invention. The DNA sequences are recorded on a computer readable
 CC medium, preferably selected from a floppy or hard disk, random access
 CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
 CC the S. aureus DNA sequences allows putative functions to be assigned so
 CC that protein-encoding or regulatory regions of commercial, therapeutic or
 CC industrial importance can be obtained. Specifically, sequences which are
 CC likely to encode antigens have been identified and these polypeptides can
 CC be used in a vaccine composition against S. aureus infection. The
 CC polypeptides can also be used in a kit for the immunodetection of
 CC S. aureus in a sample. S. aureus is implicated in numerous human diseases,
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used
 CC for recombinant production of the polypeptides. The new DNA sequences
 CC (and their fragments) are useful as primers or probes for isolating
 CC homologues of any of the S. aureus DNA sequences contained on the
 CC computer readable medium.
 SQ Sequence 9280 BP; 2958 A; 1649 C; 1309 G; 3060 T;
 Query Match 14.5%; Score 178.4; DB 1; Length 9280;
 Best Local Similarity 50.6%; Pred. No. 5.5e-29;
 Matches 484; Conservative 0; Mismatches 466; Indels 6; Gaps 2;
 QY 151 CATGAGTCTATCATCAGAGTCCGAGAGGCTCGGATCATCAGAGCGGATTC 210
 DB 9280 CTTACTGTTTGTGGAACAAATCGGCTTACAGATCATCTTGAAGATGATGATAC 9221
 QY 211 AAGCGCGAGCGCGCACTGCTGCGCAACCGCGGACAGGTGTGGCGCGAGCTGATTA 270
 DB 9220 AAGAGAGAGGTGCTGAGATGTTGCTGCAACAGCAAAAGTTTG--GATGTGATATG 9164
 QY 271 TTGCTCACTCAAGAAAGCAATGAGGATGAGGCGCGCTGCGAGCGAGGAGATC 330
 DB 9163 GTTATTAAGTTAAAGAACCTTAATCTGATATTCATATTTAAAGAGGCTTGT 9104
 QY 331 TTGTTACGTTCTTGGATTTGGCGCGTACGCTGCTGACCGATGCTGTTGATTC 390
 DB 9103 TTATTCATCTTACTTATTCAGCAATGAGAAATTAACCAAGCTTTGATAGTA 9044
 QY 391 GGCACCGTCAATGCTTTCAGAGACCGTTCAGACCGCGGAGCGGACCTACCTGCTT 450

Db 9043 AAGATATAGTATGATGACAGTACAGTACAGACCGATCTTACCATTTGTA 8984
 QY 451 GCCCGATGACGAGATCGCCGCTGACACTGCGCCGCGGCTTACACCATGATG 510
 Db 8983 TCACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8924
 QY 511 CGAACCCAGAGGGGCGCGCTGCTGATGAGGCGGCGGCTGCGAGCCGCGAC 570
 Db 8923 AACTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8864
 QY 571 GTCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 630
 Db 8863 GTACTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8804
 QY 631 GCGCGGACCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 690
 Db 8803 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8744
 QY 691 TTCTGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 730
 Db 8743 TTCTGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 8684
 QY 751 CCGTCCGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 810
 Db 8683 CAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8634
 QY 811 TCGATTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 870
 Db 8623 ACAGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8564
 QY 871 CAGGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 927
 Db 8563 CAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8504
 QY 928 GTGACGACAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 987
 Db 8503 AAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8444
 QY 988 ACCTAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1047
 Db 8443 ACCTAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 8384
 QY 1048 CCGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1103
 Db 8383 AGAGAGCATTTAAATCAATCAACCATTTATGATGATGATGATGATGATG 8328
 RESULT 8
 ID T67971 standard; DNA; 1074 BP.
 AC T67971;
 DT 15-JUL-1997 (first entry)
 DE H. pylori membrane protein ORF 05CP20518orf1.
 KW Vaccine; prevention; treatment; infection; identification;
 KW binding compound; bacterium; life cycle; activator; bacteria;
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis;
 KW membrane; amino acid; metabolism; ds.
 OS Helicobacter pylori.
 FH Key Location/Qualifiers
 FT 1.1074
 FT cds
 FT /*tag= a
 FT /note= "no stop codon given"
 PN WO9640893-A1.
 PD 19-DEC-1996.
 PE 06-JUN-1996; US-09122.
 PR 07-JUN-1995; US-487032.
 PR 01-APR-1996; US-630405.
 PA (ASTR) ASTRA AB
 PI Berglund OT, Smith D, Mellgaard BL;
 DR P-PSDB: W20718.
 PT Helicobacter pylori nucleic acid sequences and related
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori

PF Infection, and to detect Helicobacter
 PS Claim 27; Page 819; 1481p; English.
 CC The present sequence encodes a Helicobacter pylori membrane
 CC protein likely to contain four membrane spanning regions.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds,
 CC useful as potential H. pylori life cycle activators or inhibitors.
 CC The genomic sequence of H. pylori (ATCC 55679) was determined from
 CC overlapping contigs generated by mechanically shearing the
 CC bacterial DNA. The sequences were analysed for ORF of at least 180
 CC nucleotides, and the predicted coding regions defined by computer
 CC evaluation. To identify likely H. pylori antigens for vaccine
 CC development, the amino acid sequences predicted from various ORF
 CC were analysed for significant homology to other known or exported
 CC membrane proteins. Having identified and determined the sequences
 CC of interest, particular regions can be isolated from H. pylori by
 CC PCR amplification for recombinant polypeptide production, e.g. in
 CC E. coli hosts.
 SQ Sequence 1074 BP; 302 A; 189 C; 297 G; 286 T;

Query Match 7.7%; Score 94.8; DB 1; Length 1074;
 Best Local Similarity 46.1%; Pred. No. 1.2e-11;
 Matches 370; Conservative 0; Mismatches 417; Indels 15; Gaps 1;

QY 265 GATTTATTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 324
 Db 133 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 192
 QY 325 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 384
 Db 193 GCGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 252
 QY 385 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 444
 Db 253 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 312
 QY 445 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 496
 Db 313 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 372
 QY 496 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 432
 Db 373 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 432
 QY 550 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 609
 Db 433 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 492
 QY 610 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 669
 Db 493 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 552
 QY 670 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 729
 Db 553 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 612
 QY 730 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 789
 Db 613 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 672
 QY 790 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 849
 Db 673 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 732
 QY 850 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 909
 Db 733 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 792
 QY 910 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 969
 Db 793 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 852

RESULT 11

Q21833/c
ID Q21833 standard; DNA: 9890 BP.
AC Q21833;
DE 08-JUN-1992 (first entry)
DE Randomising oligonucleotide used in SPERT mRNA prep.
KM Systematic polypeptide evolution by reverse translation; SPERT;
KM ligand binding; ss.
OS Synthetic.
PN WO9202536-A.
PD 20-FEB-1992.
PE 01-AUG-1991; U05463.
PR 02-AUG-1990; US-561968.
PS (COLS) UNIV OF COLORADO.
PI Gold L, Tuerc C;
DR WPI; 92-080018/10.
PT New method of systematic polypeptide evolution by reverse
PT translation - by linking each polypeptide in sample mixt. to
PT individualised mRNA allowing further synthesis of selected
PT polypeptide(s)
PS Example; Page 55; 102pp; English.
CC The sequence is that of an example randomising oligonucleotide which
CC is used in the prepn. of mRNA encoding candidate polypeptides for the
CC method of systematic polypeptide evolution by reverse translation
CC (SPERT). The method provides a rapid way of isolating and identifying
CC polypeptide ligands which bind to target mols. The polypeptide ligands
CC can be used in e.g. assay methods, diagnostic procedures, cell sorting,
CC as activators or inhibitors of target mol. function, as probes, as
CC sequestering agents, drug delivery vehicles, modifiers of hormone
CC action and as catalysts. See also Q21830-Q21832.
SQ Sequence 390 BP; 125 A; 126 C; 133 G; 6 T;

Query Match

Best Local Similarity 4.1%; Score 50.8; DB 1; Length 390;
Matches 169; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 426 CGCCGAGGCGGACACTACCCCTGCTGCCCCGATGAGCGAAGTCCGGTGCAGTCCGCC 485
DB 366 CGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 307
QY 486 CCAGTTGGGCGTTACCACTGATGCAACCAAGGGGCGCGGTGCTGATGGCGG 545
DB 306 CGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 247
QY 546 GGTGCCCGGCTGCAACCGCGGAGCTGCTGATGCGCGCGGACCGCGGCTACAA 605
DB 246 CGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 187
QY 606 CGCAGCCCGATGCGCAACGCGGATGGGCGGACCGTTACGTTAGACATCAATCA 665
DB 186 CGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 127
QY 666 CAACATTGGCACTGACGCGGAGTCTGCGCGCGGATCCACACTGCTACTCATCGG 725
DB 126 CGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 67
QY 726 CTACGAGCTGAGAGGCGCTCAACGCTGCGGACCTGCTATTGGGCGCGTCTGCTGCC 785
DB 66 CGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTCA 7
QY 786 AGGCGC 791
DB 6 TGGCCC 1

RESULT 12
Q36859/c
ID Q36859 standard; DNA: 390 BP.
AC Q36859;
DE 22-JUN-1993 (first entry)

DE PCR primer for 5' fixed sequence contg. T7 promoter and RBS.
KM Systematic peptide evolution by reverse translation; SPERT; ligand;
KM specific; inhibitors; probes; assay; cell sorting; ss.

OS Synthetic.
PN WO9303172-A.
PD 18-FEB-1993.
PE 31-JAN-1992; U00801.
PR 01-AUG-1991; US-739055.
PS (UYRE-) UNIV RES CORP.
PI Gold L, Prihnow D, Smith JD, Tuerc C;
DR WPI; 93-076529/09.
PT Systematic polypeptide evolution by reverse translation - used
PT for prodn. of polypeptide ligand specific for desired target
PT molecule
PS Example 1; Page 84; 98pp; English.
CC SPERT is used to select novel polypeptides that bind the antibody
CC CC which are the fl progeny of a cross of NZB and NZW parents
CC (Portanova et al., J. Immunol. 144, 4633, 1990). The known epitope
CC consists of ca. 10 amino acids at the N-terminus of the histone H2B
CC protein. To make mRNA encoding candidate polypeptides a 5' fixed
CC sequence composed of a T7 promoter sequence and a ribosome binding
CC site which is recognised by both prokaryotic and eukaryotic ribosomes,
CC terminating in a restriction endonuclease site is synthesised and cloned
CC using a number of oligonucleotides (example shown). A 3' fixed sequence
CC is placed into a restriction site to provide an mRNA encoding the C-
CC terminal trailer sequence of ca. 100 nucleotides lacking stop codons.
CC In addition, a 3' primer annealing site is provided so that cDNA
CC synthesis can be accomplished on the mRNA recovered from partitioned
CC ribosome complexes. See also Q36845-63.
SQ Sequence 390 BP; 125 A; 126 C; 133 G; 6 T;

Query Match

Best Local Similarity 4.1%; Score 50.8; DB 1; Length 390;
Matches 169; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 426 CGCCGAGGCGGACACTACCCCTGCTGCCCCGATGAGCGAAGTCCGGTGCAGTCCGCC 485
DB 366 CGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 307
QY 486 CCAGTTGGGCGTTACCACTGATGCAACCAAGGGGCGCGGTGCTGATGGCGG 545
DB 306 CGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 247
QY 546 GGTGCCCGGCTGCAACCGCGGAGCTGCTGATGCGCGCGGACCGCGGCTACAA 605
DB 246 CGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 187
QY 606 CGCAGCCCGATGCGCAACGCGGATGGGCGGACCGTTACGTTAGACATCAATCA 665
DB 186 CGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 127
QY 666 CAACATTGGCACTGACGCGGAGTCTGCGCGCGGATCCACACTGCTACTCATCGG 725
DB 126 CGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 67
QY 726 CTACGAGCTGAGAGGCTGCGCAACGCTGCGGACCTGCTATTGGGCGCGTCTGCTGCC 785
DB 66 CGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTCA 7
QY 786 AGGCGC 791
DB 6 TGGCCC 1

RESULT 13

X34205
ID X34205 standard; DNA: 1413 BP.
AC X34205;
DE 06-JUL-1999 (first entry)
DE Mycobacterium species nucleic acid sequence 415.
KM Secreted protein; Mycobacterium; primer; PCR; amplification; probe;

RESULT 15

ID T68715 standard; DNA: 15872 BP.

AC T68715;

DE 01-SEP-1997 (first entry)

DE Streptomyces venezuelae polyketide synthase vep ORF1.

KM Polyketide synthase; polyhydroxyalkanoate monomer synthase;

KM polyhydroxybutyrate; biodegradable polymer; vep gene;

KM metabolic engineering; ss.

OS Streptomyces venezuelae.

FH Key Location/Qualifiers

FT cds 20..13912

FT cds /*tag- a

FT cds /*tag- b 14056..14136

FT cds /*tag- c 14148..15827

FT cds /*tag- c

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

FT cds

SQ Sequence 15872 BP; 2093 A; 6313 C; 5503 G; 1963 T;

Query Match 18; Score 50.4; DB 1; Length 15872;

Best Local Similarity 47.78; Pred. NO. 0.025; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

QY 768 TGGGCCGTCCTGTCGACAGGCGCCAGGCAC 799
DB 1900 CAGCGCCGCCACAGTCGCGGGGTCCTGTGC 1931Search completed: June 22, 2000, 15:08:21
Job time: 17827 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 12:07:29 ; Search time 5541.94 Seconds
(without alignments)
898.126 Million cell updates/sec

Title: US-09-362-485-5
Sequence: 1228
Sequence: 1 ATCTTCGAGATTAATCGAAC.....GGGAGTAAGGAAAGCATGA 1228

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4857316 seqs, 2026611550 residues
Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

- 1: em_est1:*
- 2: em_est2:*
- 3: em_est3:*
- 4: em_est4:*
- 5: em_est5:*
- 6: em_est6:*
- 7: em_est7:*
- 8: em_est8:*
- 9: em_est9:*
- 10: em_est10:*
- 11: em_est11:*
- 12: em_est12:*
- 13: em_est13:*
- 14: em_est14:*
- 15: em_est15:*
- 16: em_est16:*
- 17: em_est17:*
- 18: em_est18:*
- 19: em_est19:*
- 20: gb_est1:*
- 21: gb_est2:*
- 22: gb_est3:*
- 23: gb_est4:*
- 24: gb_est5:*
- 25: gb_est6:*
- 26: gb_est7:*
- 27: gb_est8:*
- 28: gb_est9:*
- 29: gb_est10:*
- 30: gb_est11:*
- 31: gb_est12:*
- 32: gb_est13:*
- 33: gb_est14:*
- 34: gb_est15:*
- 35: gb_est16:*
- 36: gb_est17:*
- 37: gb_est18:*
- 38: gb_est19:*
- 39: gb_est20:*
- 40: gb_est21:*
- 41: gb_est22:*
- 42: gb_est23:*
- 43: gb_est24:*
- 44: gb_est25:*

- 45: gb_est26:*
- 46: gb_est27:*
- 47: gb_est28:*
- 48: gb_est29:*
- 49: gb_est30:*
- 50: gb_est31:*
- 51: gb_est32:*
- 52: em_est20:*
- 53: em_est21:*
- 54: em_est22:*
- 55: em_est23:*
- 56: em_est24:*
- 57: em_est25:*
- 58: em_est26:*
- 59: gb_est34:*
- 60: gb_est35:*
- 61: gb_est36:*
- 62: gb_est37:*
- 63: gb_est38:*
- 64: gb_est39:*
- 65: em_est27:*
- 66: em_est28:*
- 67: em_est29:*
- 68: em_est30:*
- 69: gb_est39:*
- 70: gb_est40:*
- 71: gb_est41:*
- 72: gb_est42:*
- 73: gb_est43:*
- 74: gb_est44:*
- 75: em_est31:*
- 76: em_est32:*
- 77: em_est33:*
- 78: em_est34:*
- 79: gb_est45:*
- 80: gb_est46:*
- 81: gb_est47:*
- 82: gb_gss1:*
- 83: gb_gss2:*
- 84: gb_gss3:*
- 85: gb_gss4:*
- 86: em_gss1:*
- 87: em_gss2:*
- 88: em_gss3:*
- 89: em_gss4:*
- 90: gb_gss5:*
- 91: gb_gss6:*
- 92: gb_gss7:*
- 93: gb_gss8:*
- 94: gb_gss9:*
- 95: em_gss5:*
- 96: em_gss6:*
- 97: em_gss7:*
- 98: em_gss8:*
- 99: em_gss9:*
- 100: em_gss10:*
- 101: em_gss11:*
- 102: gb_gss10:*
- 103: gb_gss11:*
- 104: em_gss12:*
- 105: gb_gss12:*
- 106: gb_gss13:*
- 107: gb_gss14:*
- 108: gb_gss15:*
- 109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query


```

FEATURES                                pBelobAC11.
Source                                  1..846
      /organism="Drosophila melanogaster"
      /plasmid="pBelobAC11"
      /db_xref="taxon:7227"
      /clone_1lb="DrosBAC"
      /clone="BACN04N13"
      /note="end : 17"

BASE COUNT      300 a 7 c 40 g 175 t 284 others
ORIGIN

Query Match      3.8%; Score 46.6; DB 82; Length 846;
Best Local Similarity 13.0%; Pred. No. 6.7;
Matches 33; Conservative 127; Mismatches 93; Indels 0; Gaps 0.

OY 363 TCGTTCACCGATGCGGTGATTGATCCGGACACACGTCATATGCTACGAGACGCTCCA 422
      |||::: :::: :::: ||:: ||:: :::: :::: :::: :::: ::::
DB 842 TACTCTSSSTSSCTSSSSSBTBTSTTSTSTASSSSSSTSSSTSSSTSSSTSBTBT 783
      :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
OY 423 GACCCCGCAGCGCGACATCCCTGCTGCTCCCGAATGACGACGATGCGCGGACGACTCGC 482
      :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
DB 782 SSSSTTTSTTSSSTTSTTSSSTTSTGCTSSCSCSSSSSTTS:SASASJSTSTSTG:SBSS 723
      :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
OY 483 CGCCACGATGCGCGCTTACCACTGATGCGAACCAGAGGGGCCGCGGTGCTGATGCG 542
      :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
DB 722 SBTGCTSTSTBTBTBTBTSSBSBSSSSSSTTBTBTSSBSBSBTSTSTSTSS 663
      :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
OY 543 CGGGGCGCCCGCGCTGCAACCGCGCAGCTCGTGATCGGCGCGCGGACCGCGCGTAA 602
      :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
DB 662 SSSSBTSTSSBSSTSTTSTTSSSCSSSSSTSTSSBSBSBSBTSCTSSSSSCT 603
      :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
OY 603 CAACGCAAGCCCGC 615
      ||| ||| |
DB 602 CGCGSCGCCGCC 590

RESULT 7
CNS0181E/c      GSS      26-JUL-1999
LOCUS      Drosophila melanogaster genome survey sequence SP6 end of BAC
DEFINITION      BACN37H05 of DrosBAC library from Drosophila melanogaster (fruit
      fly), genomic survey sequence.
ACCESSION      AL108764
VERSION      AL108764.1 GI:5629068
KEYWORDS      GSS.
SOURCE      Drosophila melanogaster
      fruit fly.
ORGANISM      Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
      Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
      Muscomorpha; Ephydroidea; Drosophila.
REFERENCE      1 (bases 1 to 1101)
AUTHORS      Genoscope.
TITLE      Direct Submission
JOURNAL      Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage
      BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
      - Web : www.genoscope.cns.fr)
COMMENT      Determination of this BAC-end sequence was carried out as part of a
      collaboration with the European Drosophila Genome Project (EDGP) +
      http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
      library (Dros BAC) was made by Alain Billaud at CERH (Centre
      d'Etude du Polymorphisme Humain) with funding provided by a MRC
      project grant. The DNA was prepared from embryos by Alain Bucheton
      and Genevieve Payan. It has been constructed in the vector
      pBelobAC11.

FEATURES
Source      Location/Qualifiers
      1..1101
      /organism="Drosophila melanogaster"
      /plasmid="pBelobAC11"
      /db_xref="taxon:7227"
      /clone_1lb="DrosBAC"
      /clone="BACN37H05"

```

[illegible]

```

RESULT      8
LOCUS       CNS006KX
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC #
ACCESSION   BAC14AN09 of RPCI-98 library from Drosophila melanogaster (fruit
VERSION     fly), genomic survey sequence.
KEYWORDS
SOURCE      AL066051.1
ORGANISM    GI:4945019
            Drosophila melanogaster
            fruit fly.
REFERENCE   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 935)
            Genoscope.
            Direct Submission
            Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
            determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mamoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            p1 and EST libraries. A more detailed description of the library's
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
            Location:Qualifiers
            1. 935
FEATURES
            source
            /organism="Drosophila melanogaster"
            /db_xref="taxon:7227"
            /clone_lib="RPCI-98"
            /clone="BAC14AN09"
            /note="end : T7"
BASE COUNT  257 a 170 c 162 g 96 t 250 others

```



```
FEATURES
source      location/Qualifiers
1. 1101    /organism="Drosophila melanogaster"
          /plasmid="pbeloBAC11"
```

```
/clone_lib="605 -Endosperm cDNA library from Schmidt lab"  
/tissue_type="nucellar, embryo, and endosperm"  
/dev_stage="10-14 days post-pollination"  
/lab_host="DHS(alpha)"
```

/note="Organ: Kernel; Vector: PAD-GAL4-2'; site_1: EcoRI;
Site_2: XhoI; kernel endospem cDNA library from Schmidt
lab"

BASE COUNT 61 a 105 c 106 g 42 t 1 others

Query Match 3.7%; Score 45; DB 51; Length 315;
Best Local Similarity 50.7%; Pred. No. 12;
Matches 108; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

OY 391 GGCACCACTCATATTCCTACAGACACCGCCGAGCGCCGACTACCCCTGCTT 450
DB 58 GGCATGGCTCTCCATCCCAAGACACACCTGGTGGCCGACCAACTTAAACCCACC 117
OY 451 GCCCGATGAGGAGATGCGCGCTGATGCGCGCCGAGTTGGCGCTTACCACTGATG 510
DB 118 GCCCTAAGACGCGCGCGCGCGCGCGATGAGCGCGCGCGCGCTGCGGACCCG 177
OY 511 CGAACCCAGAGGCGCGCGCGCTGATGCGCGCGCGCGCGCGCGCGCGCGCGAC 570
DB 178 CGTGGCGCGCGCGCGCGCGCGCTGATGCGCGCGCGCGCGCGCGCGCGCGCG 237
OY 571 GTCGT 603
DB 238 GTGGCG 270

RESULT 15

CNS010RJ 846 bp DNA GSS 26-JUL-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC
DEFINITION BACN04N13 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION AL099337
VERSION AL099337.1 GI:5610948

KEYWORDS GSS
SOURCE fruit fly.

ORGANISM

Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 846)

REFERENCE
AUTHORS Direct Submission
TITLE Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

COMMENT

Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.edi.ac.uk. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billand at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

FEATURES

SOURCE

Location/Qualifiers
1..846
/organism="Drosophila melanogaster"
/plasmid="pBelobAC11"
/db_xref="taxon:7227"
/clone_lib="DrosBAC"
/clone="BACN04N13"
/note="end : 77"

BASE COUNT 300 a 47 c 40 g 175 t 284 others

ORIGIN

Query Match 3.7%; Score 45; DB 82; Length 846;
Best Local Similarity 16.2%; Pred. No. 14;
Matches 38; Conservative 109; Mismatches 88; Indels 0; Gaps 0;
OY 506 TGATCGAACCAGGCGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 565

DB 574 KGAGCGRRARSSAAGAGGGGGGGGGCGAGGSSSSSSSSSSSSSSSSSSSSSS 633
OY 566 CCGAGCTGCTGTGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 625
DB 634 SSS 693
OY 626 GCATGGCG 685
DB 694 SSS 753
OY 686 CCGAGTTGTGCGCGCGCGGATCCACACTCGCTACTCATCGCGCTAGCGCTGAGG 740
DB 754 SSASGASS 808

Search completed: June 22, 2000, 12:07:36
Job time: 10370 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 12:07:29 ; Search time 5541.94 Seconds
(without alignments)
898.126 Million cell updates/sec

Title: US-09-362-485-5
Perfect score: 1228
Sequence: 1 ATCTGCAGATTAATCGAAC.....GGGAGTAAGGAGGAGCATGA 1228

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 202661650 residues 9714632

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 1000000
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:*

1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: gb_est33:*
53: gb_est34:*
54: gb_est35:*
55: gb_est36:*
56: gb_est37:*
57: gb_est38:*
58: gb_est39:*
59: gb_est40:*
60: gb_est41:*
61: gb_est42:*
62: gb_est43:*
63: gb_est44:*
64: gb_est45:*
65: gb_est46:*
66: gb_est47:*
67: gb_est48:*
68: gb_est49:*
69: gb_est50:*
70: gb_est51:*
71: gb_est52:*
72: gb_est53:*
73: gb_est54:*
74: gb_est55:*
75: gb_est56:*
76: gb_est57:*
77: gb_est58:*
78: gb_est59:*
79: gb_est60:*
80: gb_est61:*
81: gb_est62:*
82: gb_est63:*
83: gb_est64:*
84: gb_est65:*
85: gb_est66:*
86: gb_est67:*
87: gb_est68:*
88: gb_est69:*
89: gb_est70:*
90: gb_est71:*
91: gb_est72:*
92: gb_est73:*
93: gb_est74:*
94: gb_est75:*
95: gb_est76:*
96: gb_est77:*
97: gb_est78:*
98: gb_est79:*
99: gb_est80:*
100: gb_est81:*
101: gb_est82:*
102: gb_est83:*
103: gb_est84:*
104: gb_est85:*
105: gb_est86:*
106: gb_est87:*
107: gb_est88:*
108: gb_est89:*
109: gb_est90:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES

No.	Score	Match Length	DB	ID	Description
C 1	57.4	4.7	925	82	CNS0091P
C 2	55.2	4.5	925	82	CNS0091P
C 3	51.8	4.2	535	74	AV390505
C 4	50.2	4.1	844	82	CNS0052P
C 5	49	4.0	910	82	CNS0060N
C 6	46.6	3.8	846	82	CNS0108J
C 7	46.6	3.8	846	82	CNS0181E
C 8	46.4	3.8	935	82	CNS006XK
C 9	46.2	3.8	645	82	CNS01213
C 10	45.8	3.7	935	82	CNS0106K
C 11	45.4	3.7	591	69	AM128786
C 12	45.2	3.7	744	83	CNS0172K
C 13	45.2	3.7	1101	83	CNS0175X
C 14	45	3.7	315	51	AT1712257
C 15	45	3.7	846	82	CNS0108J
C 16	44.8	3.6	932	82	CNS0077Q
C 17	44	3.6	843	82	CNS00C51
C 18	43.8	3.6	427	51	AT136678
C 19	43.8	3.6	512	74	AV396794
C 20	43.8	3.6	521	74	AV397116
C 21	43.8	3.6	526	74	AV394827
C 22	43.6	3.5	1101	83	CNS0160E
C 23	43.2	3.5	439	91	AQ0844063
C 24	43.2	3.5	798	82	CNS000A
C 25	43	3.5	413	74	AV396947
C 26	41.8	3.4	414	80	AM286084
C 27	41.8	3.4	465	80	AM287007
C 28	41.8	3.4	530	74	AV396357
C 29	41.8	3.4	932	82	CNS004NB
C 30	41.8	3.4	839	82	CNS0072Q
C 31	41.6	3.4	506	74	AM180713
C 32	41.6	3.4	870	82	CNS006E2
C 33	41.4	3.4	1101	82	CNS006XK
C 34	41.4	3.4	1201	83	CNS0148J
C 35	41.4	3.4	639	60	AI812194
C 36	41.2	3.4	774	105	AO327321
C 37	41.2	3.4	910	82	CNS0060N
C 38	41	3.3	605	46	AI461529
C 39	41	3.3	1101	83	CNS0175T
C 40	40.8	3.3	384	64	AM054773
C 41	40.8	3.3	385	21	D48746
C 42	40.8	3.3	437	74	AV395790
C 43	40.6	3.3	506	74	AV396405
C 44	40.4	3.3	574	91	AQ084186
C 45	40.2	3.3	916	83	CNS0155Q
					AL053013 Drosophila
					AL053013 Drosophila
					AV390505 AV390505
					AL056652 Drosophila
					AL056529 Drosophila
					AL099337 Drosophila
					AL108764 Drosophila
					AL056051 Drosophila
					AL101889 Drosophila
					AL066051 Drosophila
					AM128786 fe37d05.y
					AL108698 Drosophila
					AL112457 605069D09
					AL099337 Drosophila
					AL066742 Drosophila
					AL059666 Drosophila
					AT136678 sb32604.y
					AV396794 AV396794
					AV394827 AV394827
					AL107216 Drosophila
					AQ0844063 nbdb0022M
					AL055851 Drosophila
					AM286084 Lg1-261.B
					AM287007 Lg1-264.C
					AV396353 AV396353
					AL054280 Drosophila
					AL066742 Drosophila
					AM180713 Mga0863f
					AL064271 Drosophila
					AL078875 Drosophila
					AL103945 Drosophila
					AI812194 605086F10
					AO327321 nbdb0040D1
					AL065529 Drosophila
					AI461529 486018C05
					AM054773 Drosophila
					AM054773 ws59d08.x
					D48746 R1C51537A

ALIGNMENTS

RESULT	1
CNS0091b/c	
LOCUS	
DEFINITION	CNS0091P 925 bp DNA GSS 03-JUN-1999
ACCESSION	Drosophila melanogaster genome survey sequence TEN3 end of BAC #
KEYWORDS	BACR1D16 of Rpct-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
SOURCE	AL053013
ORGANISM	AL053013.1 GI:4934461
	GSS.
	fruit fly.
	Drosophila melanogaster
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE	1 (bases 1 to 925)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequenage : :
	BP 191 91006 Evry cedex - FRANCE (E-mail : sequefgenoscope.cns.fr
	- Web : www.genoscope.cns.fr)
COMMENT	Determination of this BAC-end sequence was carried out as part of a

collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.itslulfly.org/TheBDGP/Drosophila_melanogaster_BAC_library was prepared by Kazutoyo Oosawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP¹, 31 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES	source	location/Qualifiers
1..925		
/organism="Drosophila		melanogaster"
/db_xref="taxon:7227"		
/clone_lib="RPCI-98"		
/clone="BACR19D16"		
/note="end : TET3"		
BASE COUNT	120 a	61 c 61 g 172 t 511 others
ORIGIN		

Query Match 4.78; Score 57.4; DB 82; Length 925;

[illegible]

FEATURES	pbeloBAC11.	
SOURCE	1. 846	
	/organism="Drosophila melanogaster"	
	/plasmid="pbeloBAC11"	
	/db_xref="taxon:7227"	
	/clone_11b="DrosBAC"	
	/clone="BACN04N13"	
	/note="end : T7"	
BASE COUNT	300 a	47 c 40 g 175 t 284 others
ORIGIN		
Query Match	3.8%	Score 46.6; DB 82; Length 846;
Best Local Similarity	13.0%	Pred. No. 6.7;
Matches 33;	Conservative 127;	Mismatches 93; Indels 0; Gaps 0;
QY	363	TGCTTGACCGATGCGTTGGATTCCGGCACACGTCGATTCCTACGAGACCGTCCA 422
DB	842	TTCTGTTSTTSSCTSSSTSTBTSTTSTSTSSASSTSSSTSSSTSTSSSTSTBT 783
QY	423	GACCCCGACGGCGGACACACCCCTGCGCCCGATGAGCGAAGTCGCGGCGACACTGC 482
DB	762	SBSSTSTTTSSSTSTTSSSTSTSCSSSTSSSASASSTSTSTTCSSS 723
QY	483	CGCCGAGGTGCGCTTACACACCTGATCGAACCCAGAGGGCGCGGTGTGCTGATGG 542
DB	722	SBSSTGTSSTBTBTBTSSBSBSSTSSSTSTBTSSBSSTSSSTSTSS 663
QY	543	CGGGGTGCCCGCGCTCGAACCGCGGACGCTGCTGATCGCGCGGACCGCGGCTA 602
DB	662	SSBSSTSTSSSTSTCTTSTTSSSCSSSTSSSTSSBSSTSSCBTSSCSTSSCT 603
QY	603	CAACGACGCGCG 615
DB	602	CGCCSCGCCCC 590
RESULT 7		
CNS0181E/c		
LOCUS	CNS0181E	1101 bp DNA GSS 26-JUL-1999
DEFINITION	Drosophila melanogaster genome survey sequence SP6 end of BAC BACN37H05 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.	
ACCESSION	AL108764	
VERSION	AL108764.1	GI:5629068
KEYWORDS	GSS.	
SOURCE	fruit fly.	
ORGANISM	Drosophila melanogaster	
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
REFERENCE	1 (bases 1 to 1101)	
AUTHORS	Genoscope.	
TITLE	Direct Submission	
JOURNAL	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)	
COMMENT	Determination of this BAC end and sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) at http://www.edgp.ebi.ac.uk/. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pbeloBAC11.	
FEATURES	Location/Qualifiers	
source	1. 1101	
	/organism="Drosophila melanogaster"	
	/plasmid="pbeloBAC11"	
	/db_xref="taxon:7227"	
	/clone_11b="DrosBAC"	
	/clone="BACN37H05"	

BASE COUNT	61 a	215 c	301 g	235 t	289 others
ORIGIN	/note="end : SP6"				
Query Match	3.8%; Score 46.6; DB 83; Length 1101;				
Best Local Similarity	40.2%; Pred. No. 6.9;				
Matches 101; Conservative	40; Mismatches 109; Indels 1; Caps 1;				
OY	389 CCGGACACAGCTCATATTCCTACAGAGACGTCACAGCCGAGCGGAGCGGACCTACCCGCGC 448				
DB	308 CCCCCVVVVVVVGGMAAACCCCCCAVASSSSCCCCSSCCSCCGCGSCACSCSCSCG 249				
OY	449 TTGCCCCGATAGACGACATCCCGGATCGACTCGCCCGCCAGATTGGCGCTTACACACCTGA 508				
DB	248 AAGMGSGMSVGGCCACAAAGGCGGAGACCCGCGACGCCSSSCGCGSSCGSSCCSSSS 189				
OY	509 TCGCAACCCAAAGGGGCGCGCGGTGTGCTGATGAGCGGGGGTGCCTCGAACCGCGG 568				
DB	188 CGCGACCCSSSAGGGGCGCGCGCGCCACCCCGCGCGACVAGACCGCAC -CMWYGGSG 130				
OY	569 ACGTGTGTGTATCGCGCGCGCGCGCGCGGCTACACGACGACCGCCATTCGCAACGGCA 628				
DB	129 ACGCAAGGAGGAGAGAGCA 70				
OY	629 TGGGCGCGGACG 639				
DB	69 AGCGCCCAACG 59				
RESULT	8				
CNS006XK	935 bp DNA GSS 03-JUN-1999				
LOCUS	Drosophila melanogaster genome survey sequence T7 end of BAC #				
DEFINITION	BAC1.4809 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL065051.1 GI:4945019				
VERSION	GSS.				
KEYWORDS	fruit fly.				
SOURCE	Drosophila melanogaster				
ORGANISM	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 935)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY CUREY - FRANCE (E-mail : seq-ref@genoscope.cns.fr				
COMMENT	- Web : www.genoscope.cns.fr) determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoset in Peter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .				
FEATURES	Location/Qualifiers				
source	1..935 /organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone_lib="RPCI-98" /clone="BACRLAN09" /note="end : T7"				
BASE COUNT	257 a	170 c	162 g	96 t	250 others

d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBAC11.

FEATURES

source

Location/Qualifiers
1..744
/organism="Drosophila melanogaster"
/plasmid="pBelOBAC11"
/db_xref="taxon:7227"
/clone_1ib="DrosBAC"
/clone="BACN37108"
/note="end : SP6"

BASE COUNT 20 a 191 c 245 g 105 t 183 others

ORIGIN

Query Match 3.7%: Score 45.2; DB 83; Length 744;
Best Local Similarity 34.0%: Pred. No. 12;
Matches 87; Conservative 41; Mismatches 128; Indels 0; Gaps 0;

329 TCTGTACAGTTCTTGCATTGGCCGCGTACGCTGCTGACCAATGCGTTGATGATT 388
191 YCATTGCGCTTSGGCGCGCCYCCYCCYBYTYTCCSYGCTCCTCCTCCTGCTT 250
389 CCGGACACGATCAATGCTTACGACGCGTCCAGACCG:CGACGGGGGACACTGCGTGC 448
251 YCCTTCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 310
449 TTGCGCGATGACGAGTGGCGCGGCTGCGCTGCGCGCGCGAGTTGGCGCTTACCACTGA 508
311 CGCC 370
450 TGGACACCAAGAGGGGGCGGCTGCTGATGAGCGGGGGTCCCGGCTTGAACCGGCGG 568
371 CSCG 430
451 ACGT 584
431 NGGNGSGSGT 446

RESULT 13

CNS017SY

LOCUS

DEFINITION

CNS017SY 1101 bp DNA GSS 26-JUL-1999
Drosophila melanogaster genome survey sequence SP6 end of BAC
BACN37108 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Genoscope.
Submitted (25-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seque@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebl.ac.uk - This Drosophila melanogaster BAC
library (DrosBAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelOBAC11.

FEATURES

source

Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/plasmid="pBelOBAC11"

BASE COUNT 254 a 176 c 160 g 152 t 359 others

ORIGIN

Query Match 3.7%: Score 45.2; DB 83; Length 1101;
Best Local Similarity 12.6%: Pred. No. 13;
Matches 40; Conservative 157; Mismatches 121; Indels 0; Gaps 0;

420 CCAGACCG 479
755 CGAGAGGGGSGRGGGSGCG 814
480 CG 539
815 CSAVAS 874
540 GGGCGGGGCG 599
875 CSSSVASAAVAS 934
600 CTACACACCG 659
935 SVAS 994
660 CATGACAACTTGGCG 719
995 AAAAAAAS 1054
720 ATGCGCGTACGAGCTCGA 737
1055 VSAVAS 1072

RESULT 14

A1712257

LOCUS

DEFINITION

A1712257 315 bp mRNA EST 02-FEB-2000
60506PDJ9.x1 605 - Endosperm cDNA library from Schmidt lab zea mays
CDNA, mRNA sequ ce.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
On Jun 22, 1998 this sequence version replaced g1:3246668.
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 605069 row: D column: 09.

FEATURES

source

Location/Qualifiers
1..315
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_1ib="605 - Endosperm cDNA library from Schmidt lab"
/tissue_type="nucellar, embryo, and endosperm"
/dev_stage="10-14 days post-pollination"
/lab_host="DH5(alpha)"

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 14:53:39 ; Search time 8627.09 Seconds
(with out alignments)
-139.259 Million cell updates/sec

Title: US-09-362-485-6
Perfect score: 1235
Sequence: 1 ATCTGCAGATTAATCGAAC.....GGGAGCGATGATCGGCC 1235

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 862769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

GenBank:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_om:*
5: gb_om:*
6: gb_ph:*
7: gb_ph:*
8: gb_ph:*
9: gb_ph:*
10: gb_ph:*
11: gb_ph:*
12: gb_ph:*
13: gb_ph:*
14: gb_ph:*
15: gb_ph:*
16: gb_ph:*
17: gb_ph:*
18: gb_ph:*
19: gb_ph:*
20: gb_ph:*
21: gb_ph:*
22: gb_ph:*
23: gb_ph:*
24: gb_ph:*
25: gb_ph:*
26: gb_ph:*
27: gb_ph:*
28: gb_ph:*
29: gb_ph:*
30: gb_ph:*
31: gb_ph:*
32: gb_ph:*
33: gb_ph:*
34: gb_ph:*
35: gb_ph:*
36: gb_ph:*
37: gb_ph:*
38: gb_ph:*
39: gb_ph:*
40: gb_ph:*
41: gb_ph:*
42: gb_ph:*
43: gb_ph:*
44: gb_ph:*
45: gb_ph:*

45: gb_ph:*
46: gb_ph:*
47: gb_ph:*
48: gb_ph:*
49: gb_ph:*
50: gb_ph:*
51: gb_ph:*
52: gb_ph:*
53: gb_ph:*
54: gb_ph:*
55: gb_ph:*
56: gb_ph:*
57: gb_ph:*
58: gb_ph:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1234	99.9	1235	5	A87609 Sequence 6
2	1234	99.9	1235	5	A89749 Sequence 6
3	1232.4	99.8	1235	5	A87611 Sequence 8
4	1232.4	99.8	1235	5	A89751 Sequence 8
5	1225.4	99.2	1229	5	A87610 Sequence 7
6	1225.4	99.2	1229	5	A89750 Sequence 7
7	1223	99.0	1236	5	A87613 Sequence 10
8	1223	99.0	1236	5	A89753 Sequence 10
9	1221.4	98.9	1237	5	A87607 Sequence 4
10	1221.4	98.9	1237	5	A89747 Sequence 4
11	1221.4	98.9	56414	1	MTU002
12	1219	98.7	1235	5	A87606 Sequence 3
13	1219	98.7	1235	5	A89746 Sequence 3
14	1213.4	98.3	1228	5	A87608 Sequence 5
15	1213.4	98.3	1228	5	A89748 Sequence 5
16	1207	97.7	1245	5	A87605 Sequence 2
17	1207	97.7	1245	5	A89745 Sequence 2
18	1207	97.7	1260	5	A87604 Sequence 1
19	1207	97.7	1260	5	A89744 Sequence 1
20	1207	97.7	2584	1	MTALADH
21	1195	96.8	1209	5	A87612 Sequence 9
22	1188	96.2	1208	5	A89752 Sequence 9
23	1182	95.7	1194	2	MTU92472
24	686	55.5	720	5	A89759
25	526.4	42.6	40743	1	SC151
26	455.6	36.9	682	5	A89758
27	455.6	36.9	682	5	A89758
28	365	29.6	8046	2	AF049107
29	340.8	27.6	11514	2	AE002028
30	319.4	25.9	2780	1	RLE238118
31	318.8	25.8	1585	2	AF070716
32	302.4	24.5	1736	1	BACALDHC
33	281	22.8	1477	2	AF070715
34	278	22.5	2113	5	E01847
35	266.8	21.6	1125	5	E01847
36	249.2	20.2	137740	1	D90900
37	232.8	18.9	1375	1	BACALDHC
38	231	18.7	1776	1	BACALDHC
39	225.2	18.2	1709	1	BACALDHC
40	225.2	18.2	1709	1	BACALDHC
41	225.2	18.2	1709	1	BACALDHC
42	207.8	16.8	1714	2	AF070714
43	100	8.1	10552	2	AE001565
44	92	7.4	10512	2	AE000639
45	74	6.0	3160	1	RR005294

ALIGNMENTS

```

RESULT 1
LOCUS A87609 1235 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 6 from Patent WO9836089.
ACCESSION A87609
VERSION A87609.1 GI:6736249
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 1235)
Flohe, L. and Singh, M.
TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL Patent: WO 9836089-A 20-AUG-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
Source
1..1235
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 235 a 395 c 384 g 220 t 1 others
ORIGIN
Query Match 99.9%; Score 1234; DB 5; Length 1235;
Best Local Similarity 100.0%; Pred. No. 6.8e-167;
Matches 1235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTGCAGATTATGCAACTTTCTTCACTGAGACGTACAGTATCGAGGGGTAATC 60
DB 1 ATCTGCAGATTATGCAACTTTCTTCACTGAGACGTACAGTATCGAGGGGTAATC 60
QY 61 ATGCGCGTGTGATTTCGACGAGACCAAAAAACAAGATTCCGGGTGCCATCACCCG 120
DB 61 ATGCGCGTGTGATTTCGACGAGACCAAAAAACAAGATTCCGGGTGCCATCACCCG 120
QY 121 GCCGGCGTGGGGAAGTAAACCGGTGCGCATGAGTGTCATCCAGCAGAGTCCGGA 180
DB 121 GCCGGCGTGGGGAAGTAAACCGGTGCGCATGAGTGTCATCCAGCAGAGTCCGGA 180
QY 121 GCCGGCGTGGGGAAGTAAACCGGTGCGCATGAGTGTCATCCAGCAGAGTCCGGA 180
DB 121 GCCGGCGTGGGGAAGTAAACCGGTGCGCATGAGTGTCATCCAGCAGAGTCCGGA 180
QY 181 GAGGCGTGGGTATACGACGACGCGGATTTCAAGCGGCGGCAACTGTTGGGACG 240
DB 181 GAGGCGTGGGTATACGACGACGCGGATTTCAAGCGGCGGCAACTGTTGGGACG 240
QY 241 GCCGACAGAGTGTGGGCGGACGCTATTATTCGTCACAGTCAAGACCCATACGGCG 300
DB 241 GCCGACAGAGTGTGGGCGGACGCTATTATTCGTCACAGTCAAGACCCATACGGCG 300
QY 301 GAATAGGCGGCGCTGCGACACGCGGATTTGTACGTTCTTGATTTGGCGCGTCAC 360
DB 301 GAATAGGCGGCGCTGCGACACGCGGATTTGTACGTTCTTGATTTGGCGCGTCAC 360
QY 361 GTGCTTGACGAGTGGGTTGTGATTCGCGACACGTCATTTGCCATACGACCGTCC 420
DB 361 GTGCTTGACGAGTGGGTTGTGATTCGCGACACGTCATTTGCCATACGACCGTCC 420
QY 421 AAGCGCGGACGCGGACGACGCGGCTGTCGCGGATGAGCAAGTGGCGGCTGCACTG 480
DB 421 AAGCGCGGACGCGGACGACGCGGCTGTCGCGGATGAGCAAGTGGCGGCTGCACTG 480
QY 481 CCGCCAGAGTGGCGCTTACCACTGATGCAACCAAGGGGCGCGGTGTGTGATGG 540
DB 481 CCGCCAGAGTGGCGCTTACCACTGATGCAACCAAGGGGCGCGGTGTGTGATGG 540
QY 541 GCGGGGTGCGCGGCTGCAACGCGGCGAGTGTGTGATTCGCGGCGCGGCGGCT 600
DB 541 GCGGGGTGCGCGGCTGCAACGCGGCGAGTGTGTGATTCGCGGCGCGGCGGCT 600
QY 601 ACAAGCGAGCGGACGACGAGTGGGCGGCGGCGGCTTACGAGTTAGACATCAACA 660
DB 601 ACAAGCGAGCGGACGACGAGTGGGCGGCGGCGGCGGCTTACGAGTTAGACATCAACA 660
QY 661 TCGACAAACTTGGCAACTGACGCGGAGTTGTGCGGCGGATCCACACTGCTACTAT 720

```

```

DB 661 TCGACAAACTTGGCAACTGACGCGGAGTTGTGCGGCGGATCCACACTGCTACTAT 720
QY 721 CGGCGTACAGAGTCCAGAGGTGCGGTCAAAAGTCCGACCTGGTGTGGGCGGCTCGG 780
DB 721 CGGCGTACAGAGTCCAGAGGTGCGGTCAAAAGTCCGACCTGGTGTGGGCGGCTCGG 780
QY 781 TCGCAGGCGCCAAAGGACCCCAATTAGTCTCGAATTCATTTGCGGATAGAAACAG 840
DB 781 TCGCAGGCGCCAAAGGACCCCAATTAGTCTCGAATTCATTTGCGGATAGAAACAG 840
QY 841 GTGCGGTACTGTGATATAGCATATGACATGACAGGCGGCGGTGTTCGAAGCTCAGACCGA 900
DB 841 GTGCGGTACTGTGATATAGCATATGACATGACAGGCGGCGGTGTTCGAAGCTCAGACCGA 900
QY 901 CCACCTACGACACCCGCGAGTTCGCGGTGACAGACGCGTGTACTGCGTGGGGAACA 960
DB 901 CCACCTACGACACCCGCGAGTTCGCGGTGACAGACGCGTGTACTGCGTGGGGAACA 960
QY 961 TGGCGCGCTGGTGGCGGAAGAGTGCAGCTACGCGGTGACCAACGCGAGATGCCGTATG 1020
DB 961 TGGCGCGCTGGTGGCGGAAGAGTGCAGCTACGCGGTGACCAACGCGAGATGCCGTATG 1020
QY 1021 TGGTGGAGCTTGGCGACATAGGCTGGGCGGCGGTGCGGTGCAATCCGGCACTAGCCA 1080
DB 1021 TGGTGGAGCTTGGCGACATAGGCTGGGCGGCGGTGCGGTGCAATCCGGCACTAGCCA 1080
QY 1081 AAGGCTTTTCGACGACGAAAGGCGGCTTACTGTCCACAGGCGTGGCGACGCTGGGGG 1140
DB 1081 AAGGCTTTTCGACGACGAAAGGCGGCTTACTGTCCACAGGCGTGGCGACGCTGGGGG 1140
QY 1141 TGGCGTTACAGGCGCGCGGCGGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 1200
DB 1141 TGGCGTTACAGGCGCGCGGCGGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 1200
QY 1201 ACAAGTGGGAGTAAAGGAAAGCATGATGTGCGCC 1235
DB 1201 ACAAGTGGGAGTAAAGGAAAGCATGATGTGCGCC 1235

RESULT 2
LOCUS A89749 1235 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 6 from Patent WO9832862.
ACCESSION A89749
VERSION A89749.1 GI:6738283
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 1235)
Flohe, L. and Singh, M.
TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARITIMUM
JOURNAL Patent: WO 9832862-A 30-JUL-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
Source
1..1235
/organism="unidentified"
/db_xref="taxon:32544"
BASE COUNT 235 a 395 c 384 g 220 t 1 others
ORIGIN
Query Match 99.9%; Score 1234; DB 5; Length 1235;
Best Local Similarity 100.0%; Pred. No. 6.8e-167;
Matches 1235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTGCAGATTATGCAACTTTCTTCACTGAGACGTACAGTATCGAGGGGTAATC 60
DB 1 ATCTGCAGATTATGCAACTTTCTTCACTGAGACGTACAGTATCGAGGGGTAATC 60
QY 61 ATGCGCGTGTGATTTCGACGAGACCAAAAAACAAGATTCCGGGTGCCATCACCCG 120

```

```

Db 61 ATGCGGTGCTATTCGACCGAGCAAAAACAAGATTCGGGTGGCATCACCCG 120
Qy 121 GCGGGGTGCGGAATTAACCCGTGTCGTCATGAGTCTCATCGAGCAGTGC 180
Qy 121 GCGGGGTGCGGAATTAACCCGTGTCGTCATGAGTCTCATCGAGCAGTGC 180
Db 121 GCGGGGTGCGGAATTAACCCGTGTCGTCATGAGTCTCATCGAGCAGTGC 180
Qy 181 GAGGGGTGCGATATCAACGAGCGGATTCAGAGGGCGAGCGCCAACTGTGGCACC 240
Db 181 GAGGGGTGCGATATCAACGAGCGGATTCAGAGGGCGAGCGCCAACTGTGGCACC 240
Db 181 GAGGGGTGCGATATCAACGAGCGGATTCAGAGGGCGAGCGCCAACTGTGGCACC 240
Qy 241 GCGGACGAGTGTGGCGCAGCTGATTTATGCTCAAGGTCAAGAACCGATAGCGCG 300
Qy 241 GCGGACGAGTGTGGCGCAGCTGATTTATGCTCAAGGTCAAGAACCGATAGCGCG 300
Db 241 GCGGACGAGTGTGGCGCAGCTGATTTATGCTCAAGGTCAAGAACCGATAGCGCG 300
Qy 301 GAATACGCGCGCTGCGACACAGGGCGATCTGTCTACGCTTCTTGGATTGGCGGTCAC 360
Qy 301 GAATACGCGCGCTGCGACACAGGGCGATCTGTCTACGCTTCTTGGATTGGCGGTCAC 360
Db 301 GAATACGCGCGCTGCGACACAGGGCGATCTGTCTACGCTTCTTGGATTGGCGGTCAC 360
Qy 361 GTGCTTGACACGATGCGTGTGATTCGCGCACACGTCATTTGCCCTACGAGACCGTCC 420
Qy 361 GTGCTTGACACGATGCGTGTGATTCGCGCACACGTCATTTGCCCTACGAGACCGTCC 420
Db 361 GTGCTTGACACGATGCGTGTGATTCGCGCACACGTCATTTGCCCTACGAGACCGTCC 420
Qy 421 AGACCGCGCAGCGGACGACTACCCCTGCTTCCCGATGAGCGAATGCCCGGTCCACTCG 480
Qy 421 AGACCGCGCAGCGGACGACTACCCCTGCTTCCCGATGAGCGAATGCCCGGTCCACTCG 480
Db 421 AGACCGCGCAGCGGACGACTACCCCTGCTTCCCGATGAGCGAATGCCCGGTCCACTCG 480
Qy 481 CCGCCCAAGTGTGGCGCTTACCATGATGAGAACCAAGGGGGCGGCTGTATG 540
Qy 481 CCGCCCAAGTGTGGCGCTTACCATGATGAGAACCAAGGGGGCGGCTGTATG 540
Db 481 CCGCCCAAGTGTGGCGCTTACCATGATGAGAACCAAGGGGGCGGCTGTATG 540
Qy 541 GCGGGGTGCGCGCTGCGAACCGCGGACGTCGTGTGATCGCGCGGACCGCGGCT 600
Qy 541 GCGGGGTGCGCGCTGCGAACCGCGGACGTCGTGTGATCGCGCGGACCGCGGCT 600
Db 541 GCGGGGTGCGCGCTGCGAACCGCGGACGTCGTGTGATCGCGCGGACCGCGGCT 600
Qy 601 ACAACGACGCGCGCATGCGCAACGCGATGGCGCGACCGCTTAGCGTTCTAGAAATACA 660
Qy 601 ACAACGACGCGCGCATGCGCAACGCGATGGCGCGACCGCTTAGCGTTCTAGAAATACA 660
Db 601 ACAACGACGCGCGCATGCGCAACGCGATGGCGCGACCGCTTAGCGTTCTAGAAATACA 660
Qy 661 TCGACAAACTTGGGCAATGAGACGCGGAGTTGCGCGCGGATCGACACGCTACTCAT 720
Qy 661 TCGACAAACTTGGGCAATGAGACGCGGAGTTGCGCGCGGATCGACACGCTACTCAT 720
Db 661 TCGACAAACTTGGGCAATGAGACGCGGAGTTGCGCGCGGATCGACACGCTACTCAT 720
Qy 721 CCGGCTACGAGTCTGAGAGTGGCGTCAAAAGTCCGACGCTGTGATTTGGCGGCTCTCG 780
Qy 721 CCGGCTACGAGTCTGAGAGTGGCGTCAAAAGTCCGACGCTGTGATTTGGCGGCTCTCG 780
Db 721 CCGGCTACGAGTCTGAGAGTGGCGTCAAAAGTCCGACGCTGTGATTTGGCGGCTCTCG 780
Qy 781 TCGCAGGCGCGAAGGCAACCAATTAATCTGATTCGATTCGCGCATATGAACACAG 840
Qy 781 TCGCAGGCGCGAAGGCAACCAATTAATCTGATTCGATTCGCGCATATGAACACAG 840
Db 781 TCGCAGGCGCGAAGGCAACCAATTAATCTGATTCGATTCGCGCATATGAACACAG 840
Qy 841 GTGCGTACTGATGATATAGCATGACACAGGCGGCTTTTCGAAAGGCTACGACGCA 900
Qy 841 GTGCGTACTGATGATATAGCATGACACAGGCGGCTTTTCGAAAGGCTACGACGCA 900
Db 841 GTGCGTACTGATGATATAGCATGACACAGGCGGCTTTTCGAAAGGCTACGACGCA 900
Qy 901 CCACCTACGACACCGACGCTGCGCTGACACAGCCTGTTTACTGCTGGCGAACA 960
Qy 901 CCACCTACGACACCGACGCTGCGCTGACACAGCCTGTTTACTGCTGGCGAACA 960
Db 901 CCACCTACGACACCGACGCTGCGCTGACACAGCCTGTTTACTGCTGGCGAACA 960
Qy 961 TCGCGGCGTGGTGGCGAAGCGTGAAGTGTGAGTGTGACCAAGCGACGATGCGCTATG 1020
Qy 961 TCGCGGCGTGGTGGCGAAGCGTGAAGTGTGAGTGTGACCAAGCGACGATGCGCTATG 1020
Db 961 TCGCGGCGTGGTGGCGAAGCGTGAAGTGTGAGTGTGACCAAGCGACGATGCGCTATG 1020
Qy 1021 TGTCTGAGCTTGGCGACCATGCTGCGCGCGCGCTGCGCTGCAATTCGCGCACTAACCA 1080
Qy 1021 TGTCTGAGCTTGGCGACCATGCTGCGCGCGCGCTGCGCTGCAATTCGCGCACTAACCA 1080
Db 1021 TGTCTGAGCTTGGCGACCATGCTGCGCGCGCGCTGCGCTGCAATTCGCGCACTAACCA 1080
Qy 1081 AAGGCTTTCGACGACGAGAGGGGGGTTACTGTCGGAAGGGGTGGCCACCACTGGGG 1140
Qy 1081 AAGGCTTTCGACGACGAGAGGGGGGTTACTGTCGGAAGGGGTGGCCACCACTGGGG 1140
Db 1081 AAGGCTTTCGACGACGAGAGGGGGGTTACTGTCGGAAGGGGTGGCCACCACTGGGG 1140
Qy 1141 TCGCGTTACCGAGCGCGCGACGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Qy 1141 TCGCGTTACCGAGCGCGCGACGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db 1141 TCGCGTTACCGAGCGCGCGACGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200

```

```

Qy 1201 ACACGTGCGAGTAAGGGAAGCATGATGTCGCC 1235
Db 1201 ACACGTGCGAGTAAGGGAAGCATGATGTCGCC 1235

RESULT 3
LOCUS A87611 1235 bp DNA
DEFINITION Sequence 8 from Patent WO9836089.
ACCESSION A87611
VERSION A87611.1 GI:6736251
KEYWORDS
SOURCE
ORGANISM
unidentified.
unclassified.

REFERENCE
1 (bases 1 to 1235)
AUTHORS Flohe, L. and Singh, M.
TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL Patent: WO 9836089-A-20-AUG-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)

FEATURES
source
1..1235
location/Qualifiers
BASE COUNT 236 a 394 c 385 g 220 t
ORIGIN

Query Match 99.8%; Score 1232.4; DB 5; Length 1235;
Best Local Similarity 99.8%; Pred. No. 1.le-16;
Matches 1233; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATCTTGACAGATTAATGCAACTTCTTACACTGAAAGCGTACAGTATGAGAGGGTAAATC 60
Db 1 ATCTTGACAGATTAATGCAACTTCTTACACTGAAAGCGTACAGTATGAGAGGGTAAATC 60
Qy 61 ATGCGCGTGGTATTTCGACGAGACCAAAACAAGATTCGGGTGGCCATACCCCG 120
Qy 61 ATGCGCGTGGTATTTCGACGAGACCAAAACAAGATTCGGGTGGCCATACCCCG 120
Db 61 ATGCGCGTGGTATTTCGACGAGACCAAAACAAGATTCGGGTGGCCATACCCCG 120
Qy 121 GCGGCGTGGCGAATTAACCGCTGCTGGCAATGAGTGCATCCAGGAGGTGGCGGA 180
Qy 121 GCGGCGTGGCGAATTAACCGCTGCTGGCAATGAGTGCATCCAGGAGGTGGCGGA 180
Db 121 GCGGCGTGGCGAATTAACCGCTGCTGGCAATGAGTGCATCCAGGAGGTGGCGGA 180
Qy 181 GAGGGCTGGCTATACGACGCGGATTTCAAGGCGGACGCGCAACTGGTGGCAC 240
Qy 181 GAGGGCTGGCTATACGACGCGGATTTCAAGGCGGACGCGCAACTGGTGGCAC 240
Db 181 GAGGGCTGGCTATACGACGCGGATTTCAAGGCGGACGCGCAACTGGTGGCAC 240
Qy 241 GCGGACGAGTGTGGCGCGACGCTGATTTATGCTCAAGGTCAAGAACCGATAGCGCG 300
Qy 241 GCGGACGAGTGTGGCGCGACGCTGATTTATGCTCAAGGTCAAGAACCGATAGCGCG 300
Db 241 GCGGACGAGTGTGGCGCGACGCTGATTTATGCTCAAGGTCAAGAACCGATAGCGCG 300
Qy 301 GAATACGCGCGCTGGACAGCGGAGATCTGTACAGTCTTTCGATTTGGCGGCTCAC 360
Qy 301 GAATACGCGCGCTGGACAGCGGAGATCTGTACAGTCTTTCGATTTGGCGGCTCAC 360
Db 301 GAATACGCGCGCTGGACAGCGGAGATCTGTACAGTCTTTCGATTTGGCGGCTCAC 360
Qy 361 GTGCTTGACACGATGCGTGTGATTCGCGCACACGTCATTTGCCCTACGAGACCGTCC 420
Qy 361 GTGCTTGACACGATGCGTGTGATTCGCGCACACGTCATTTGCCCTACGAGACCGTCC 420
Db 361 GTGCTTGACACGATGCGTGTGATTCGCGCACACGTCATTTGCCCTACGAGACCGTCC 420
Qy 421 AGACCGCGCAGCGGACGACGCTGCTGCGCGCATGAGCGCAAGTGGCGGTGACATCG 480
Qy 421 AGACCGCGCAGCGGACGACGCTGCTGCGCGCATGAGCGCAAGTGGCGGTGACATCG 480
Db 421 AGACCGCGCAGCGGACGACGCTGCTGCGCGCATGAGCGCAAGTGGCGGTGACATCG 480
Qy 481 CCGCCCAAGTGTGGCGCTTACCATGATGAGAACCAAGGGGGCGGCTGTGCTGATGG 540
Qy 481 CCGCCCAAGTGTGGCGCTTACCATGATGAGAACCAAGGGGGCGGCTGTGCTGATGG 540
Db 481 CCGCCCAAGTGTGGCGCTTACCATGATGAGAACCAAGGGGGCGGCTGTGCTGATGG 540
Qy 541 GCGGGGTGCGCGCTGCGAACCGCGGACGTCGTGTGATCGCGCGGACCGCGGCT 600
Qy 541 GCGGGGTGCGCGCTGCGAACCGCGGACGTCGTGTGATCGCGCGGACCGCGGCT 600
Db 541 GCGGGGTGCGCGCTGCGAACCGCGGACGTCGTGTGATCGCGCGGACCGCGGCT 600

```

QY 601 ACACGACGACCCGATCGCCCAACGCGATGAGGCGGACCGCTTACGGTTCAGACATCAACA 660
| | | | |
Db 601 ACACGACGACCCGATCGCCCAACGCGATGAGGCGGACCGCTTACGGTTCAGACATCAACA 660
QY 661 TCGACAAACTTCGCACTCGACGCGGAGTTCTGCGGCGGATCCACACTCGCTACTCAT 720
| | | | |
Db 661 TCGACAAACTTCGCACTCGACGCGGAGTTCTGCGGCGGATCCACACTCGCTACTCAT 720
QY 721 CGGCGTACGAGCTGAGGAGTGCCTCAACGTCGCGACCTGCTGATGTTGGGCGGCTCTGG 780
| | | | |
Db 721 CGGCGTACGAGCTGAGGAGTGCCTCAACGTCGCGACCTGCTGATGTTGGGCGGCTCTGG 780
QY 781 TCGCAGGCGCCCAAGGACCCCAATTAAGTCTCGAATTCATTGTCGCGCATATGAAACAG 840
| | | | |
Db 781 TCGCAGGCGCCCAAGGACCCCAATTAAGTCTCGAATTCATTGTCGCGCATATGAAACAG 840
QY 841 GTGGCGTACTGGTGAATATAGCATCGACGAGGCGGCGGCTTCCAAAGGCTCAGACCGA 900
| | | | |
Db 841 GTGGCGTACTGGTGAATATAGCATCGACGAGGCGGCGGCTTCCAAAGGCTCAGACCGA 900
QY 901 CCACCTACGACCCGACCGAGTTCGCGCTGACAGACAGCTGTTTACTGCTGCGGAAACA 960
| | | | |
Db 901 CCACCTACGACCCGACCGAGTTCGCGCTGACAGACAGCTGTTTACTGCTGCGGAAACA 960
QY 961 TCGCCGCTCGGTCGCGGAAAGAGTGCAGCTACGCGCTGACCAACGCGAGATGCCGTATG 1020
| | | | |
Db 961 TCGCCGCTCGGTCGCGGAAAGAGTGCAGCTACGCGCTGACCAACGCGAGATGCCGTATG 1020
QY 1021 TGCCTGAGCTTCGCAACATGAGTGGGCGGCGGCTGCGGTCAGATCGGAGTACGCA 1080
| | | | |
Db 1021 TGCCTGAGCTTCGCAACATGAGTGGGCGGCGGCTGCGGTCAGATCGGAGTACGCA 1080
QY 1081 AAGGCTTTTCGACGACGAGGAGGCGGTTACTGTCCGAACGCGGTGCGCACGACGCTGGGG 1140
| | | | |
Db 1081 AAGGCTTTTCGACGACGAGGAGGCGGTTACTGTCCGAACGCGGTGCGCACGACGCTGGGG 1140
QY 1141 TGGCGTACGAGCGGCGGCGGAGGCGGCTGCTGACTCTCGGCGGCTGCTGCTGCTGCTG 1200
| | | | |
Db 1141 TGGCGTACGAGCGGCGGCGGAGGCGGCTGCTGACTCTCGGCGGCTGCTGCTGCTGCTG 1200
QY 1201 ACACGTCGAGATAGGAGGAGGATGATGTCGCC 1235
| | | | |
Db 1201 ACACGTCGAGATAGGAGGAGGATGATGTCGCC 1235
RESULT 4
A89751 1235 bp DNA
LOCUS A89751 Sequence 8 from Patent W09832862. PAT 22-JAN-2000
DEFINITION A89751
ACCESSION A89751
VERSION A89751.1 GI:6738285
KEYWORDS
SOURCE .
ORGANISM unidentified.
unclassified.
unclassified.
REFERENCE 1 (bases 1 to 1235)
AUTHORS Flohe L. and Singh M.
TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
JOURNAL Patent: WO 9832862-A 30-JUL-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
source 1. 1235
location/Qualifiers
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 236 a 394 c 385 g 220 t
ORIGIN

QY 1 ATCTTGACGATTAATGCAACTTCTTCACTGAGCGTACGATGAGGCGGTATC 60
| | | | |
Db 1 ATCTTGACGATTAATGCAACTTCTTCACTGAGCGTACGATGAGGCGGTATC 60
QY 61 ATGCGCGTGGTATTCGACCGAGACCAAAACAGAAATTCGGGTGGCCATACCCCG 120
| | | | |
Db 61 ATGCGCGTGGTATTCGACCGAGACCAAAACAGAAATTCGGGTGGCCATACCCCG 120
QY 121 GCGGCGTGGGGAATACCGCTGCTGGCCATGAGGTGCTATCCAGGCGAGGCGCGGA 180
| | | | |
Db 121 GCGGCGTGGGGAATACCGCTGCTGGCCATGAGGTGCTATCCAGGCGAGGCGCGGA 180
QY 181 GAGGCGTGGGATTCACCGACCGGATTTCAAGCGGCGAGGCGGCAACTGCTGGCACC 240
| | | | |
Db 181 GAGGCGTGGGATTCACCGACCGGATTTCAAGCGGCGAGGCGGCAACTGCTGGCACC 240
QY 241 GCGGACGAGGTGGGCGGAGCGGATTTATGCTCAAGGTCAAAAGAACGATAGCGCG 300
| | | | |
Db 241 GCGGACGAGGTGGGCGGAGCGGATTTATGCTCAAGGTCAAAAGAACGATAGCGCG 300
QY 301 GAATACGCGCGCTGCGACAGGCGGATCTTGTACAGTCTTGCATTTGGCGCGTAC 360
| | | | |
Db 301 GAATACGCGCGCTGCGACAGGCGGATCTTGTACAGTCTTGCATTTGGCGCGTAC 360
QY 361 GTGCTTGACCGATGCGTGTGATTCGCGACACGCTCAATTCCTTACGAGACGCTCC 420
| | | | |
Db 361 GTGCTTGACCGATGCGTGTGATTCGCGACACGCTCAATTCCTTACGAGACGCTCC 420
QY 421 AGACCGCGGAGGCGGACATACCGCTGCTGCGGATGAGGAGTGGCGGCTGATCG 480
| | | | |
Db 421 AGACCGCGGAGGCGGACATACCGCTGCTGCGGATGAGGAGTGGCGGCTGATCG 480
QY 481 CCGCCGAGGTTGGGCTTACCACTGATGCGAACCCCAAGGCGGCGGCTGCTGATG 540
| | | | |
Db 481 CCGCCGAGGTTGGGCTTACCACTGATGCGAACCCCAAGGCGGCGGCTGCTGATG 540
QY 541 GCGGCGTCCCGGCGTGAACCGGCGGAGTGTGATGCGCGCGGCGCGCGCT 600
| | | | |
Db 541 GCGGCGTCCCGGCGTGAACCGGCGGAGTGTGATGCGCGCGGCGCGCGCT 600
QY 601 ACACGACGCGCGATGCGCAAGGCGATGGGCGGCGGCTTACGCTGATGATCAACA 660
| | | | |
Db 601 ACACGACGCGCGATGCGCAAGGCGATGGGCGGCGGCTTACGCTGATGATCAACA 660
QY 661 TCGACAAACTTCGCACTCGACGCGGAGTCTCGGCGGAGTCCACACTCGCTACTCAT 720
| | | | |
Db 661 TCGACAAACTTCGCACTCGACGCGGAGTCTCGGCGGAGTCCACACTCGCTACTCAT 720
QY 721 CGGCTACGAGCTGAGGAGTGCCTCAACGTCGCACTGCTGATTTGGGCGGCTCTCG 780
| | | | |
Db 721 CGGCTACGAGCTGAGGAGTGCCTCAACGTCGCACTGCTGATTTGGGCGGCTCTCG 780
QY 781 TCGCAGGCGCCCAAGGACCCCAATTAAGTCTGATGCTGAGGCGGATATGAAACAG 840
| | | | |
Db 781 TCGCAGGCGCCCAAGGACCCCAATTAAGTCTGATGCTGAGGCGGATATGAAACAG 840
QY 841 GTGGGTAAGTGGTGAATATAGCATCGACGAGGCGGCTGTTTGAAGGCTCAGACCGA 900
| | | | |
Db 841 GTGGGTAAGTGGTGAATATAGCATCGACGAGGCGGCTGTTTGAAGGCTCAGACCGA 900
QY 901 CCACCTACGACCCGACCGAGTTCGCGCTGACAGACAGCTGTTTACTGCTGCGGAAACA 960
| | | | |
Db 901 CCACCTACGACCCGACCGAGTTCGCGCTGACAGACAGCTGTTTACTGCTGCGGAAACA 960
QY 961 TCGCCGCTCGGTCGCGGAAAGAGTGCAGCTACGCGCTGACCAACGCGAGATGCCGTATG 1020
| | | | |
Db 961 TCGCCGCTCGGTCGCGGAAAGAGTGCAGCTACGCGCTGACCAACGCGAGATGCCGTATG 1020
QY 1021 TGCCTGAGCTTCGCAACATGAGTGGGCGGCGGCTGCGGTCAGATCGGAGTACGCA 1080
| | | | |
Db 1021 TGCCTGAGCTTCGCAACATGAGTGGGCGGCGGCTGCGGTCAGATCGGAGTACGCA 1080
QY 1081 AAGGCTTTTCGACGACGAGGAGGCGTACTGTGCAAGGCGGTGCGACGACCTGGGG 1140
| | | | |

Db 1081 AAGGCTTTGACGACGACGAGGAGGCGTTACTGTCCGAAAGGGGTGGCCACGACGCTGGGG 1140
 QY 1141 TGCCGTTACCGAGAGCCCGCCAGCGCTGCTGGCTGACTCTGGCGGCTGTTACGCGGANC 1200
 Db 1141 TGCCGTTACCGAGAGCCCGCCAGCGCTGCTGGCTGACTCTGGCGGCTGTTACGCGGAGC 1200
 QY 1201 ACAGCTGGAGTAGGAGGAGGATGATGCGGC 1235
 Db 1201 ACAGCTGGAGTAGGAGGAGGATGATGCGGC 1235

RESULT 5
 A87610 1229 bp DNA PAT 22-JAN-2000
 LOCUS A87610
 DEFINITION Sequence 7 from Patent WO9836089.
 ACCESSION A87610 GI:6736250
 VERSION A87610.1 GI:6736250
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1229)
 AUTHORS Flohe, L. and Singh, M.
 TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
 JOURNAL Patent: WO 9836089-A 20-AUG-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source 1..1229
 Location/Qualifiers
 /organism="unidentified"
 /db_xref="taxon:32644"

BASE COUNT 236 a 391 c 382 g 219 t 1 others
 ORIGIN

Query Match 99.2%; Score 1225.4; DB 5; Length 1229;
 Best Local Similarity 99.8%; Pred. No. 1.1e-165;
 Matches 1226; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACTTTCAGATTATATGCACTTTCTTACACATGACGATACAGTATCGAGAGGGTATATC 60
 Db 1 ACTTTCAGATTATATGCACTTTCTTACACATGACGATACAGTATCGAGAGGGTATATC 60
 QY 61 ATGCGCGTGGTATTTCCGACGAGACCAAAACAAAGATTCGGGGTGGCCATCACCAG 120
 Db 61 ATGCGCGTGGTATTTCCGACGAGACCAAAACAAAGATTCGGGGTGGCCATCACCAG 120
 QY 121 GCGGCGTGGCGAATTAACCCGCTGCGCATGAGTGTCTATCCAGGACAGTGGCGGA 180
 Db 121 GCGGCGTGGCGAATTAACCCGCTGCGCATGAGTGTCTATCCAGGACAGTGGCGGA 180
 QY 181 GAGGCTCGGCTATCAACGACGCGGATTTCAAGGGGCGAGCGGCACTGGTGGCACC 240
 Db 181 GAGGCTCGGCTATCAACGACGCGGATTTCAAGGGGCGAGCGGCACTGGTGGCACC 240
 QY 241 GCGGACGAGTGGGCGGCGGCGCTGATTTATGCTCAAGGTCAAAAGACGATAGGGGG 300
 Db 241 GCGGACGAGTGGGCGGCGGCGCTGATTTATGCTCAAGGTCAAAAGACGATAGGGGG 300
 QY 301 GAATACGGCGCTCGACACAGGCGGATTTCTTCAAGTTCCTTCAATTTGGCGGCTCAC 360
 Db 301 GAATACGGCGCTCGACACAGGCGGATTTCTTCAAGTTCCTTCAATTTGGCGGCTCAC 360
 QY 361 GTGCTTGCACCAATGCTTTGTGATTCGCGACACAGTAAATGGCTACAGAACGCTGC 420
 Db 361 GTGCTTGCACCAATGCTTTGTGATTCGCGACACAGTAAATGGCTACAGAACGCTGC 420
 QY 421 AGACGCGGAGGCGGCTACCCCTGCTGCGGATGAGCGAAGTGGCGGCTGACTGC 480
 Db 421 AGACGCGGAGGCGGCTACCCCTGCTGCGGATGAGCGAAGTGGCGGCTGACTGC 480
 QY 481 CCGCCGAGGTTGGGCTTACCACTGATGGAACCCAGAGGGGGCGGCTGTGCTATGG 540
 Db 481 CCGCCGAGGTTGGGCTTACCACTGATGGAACCCAGAGGGGGCGGCTGTGCTATGG 540

QY 541 GCGGGGTGCCCCGCTGCAACCGGCGAGCGTGTGATCGGGGCGGCGGCGGCT 600
 Db 541 GCGGGGTGCCCCGCTGCAACCGGCGAGCGTGTGATCGGGGCGGCGGCGGCT 600
 QY 601 ACAACGAGCCCGCATCGCCCAACGGCATGGGCGGAGCCGTTACGATTAACA 660
 Db 601 ACAACGAGCCCGCATCGCCCAACGGCATGGGCGGAGCCGTTACGATTAACA 660
 QY 661 TCGCAAACTTGGCACTGAGCGCCGAGTTCGGGCGCGGATTCACACCTGCTACTCAT 720
 Db 661 TCGCAAACTTGGCACTGAGCGCCGAGTTCGGGCGCGGATTCACACCTGCTACTCAT 720
 QY 721 CGGCTTACGAGCTGAGAGGTGCGCTCAAGCGGCGGAGCTGTGATTTGGGCGCTCTCG 780
 Db 721 CGGCTTACGAGCTGAGAGGTGCGCTCAAGCGGCGGAGCTGTGATTTGGGCGCTCTCG 780
 QY 781 TGCCAGGCGCCAGGACCCCAATTAAGTTCGATTTACTTGGCGCATTTGAAACAG 840
 Db 781 TGCCAGGCGCCAGGACCCCAATTAAGTTCGATTTACTTGGCGCATTTGAAACAG 840
 QY 841 GTGGGCTACTGTGATATATACCATCGACAGCGGCGGCTGTTGAAAGGCTCACAGCGA 900
 Db 841 GTGGGCTACTGTGATATATACCATCGACAGCGGCGGCTGTTGAAAGGCTCACAGCGA 900
 QY 901 CCACCTACGACACCCGCGAGCGTTGCGCGTGCACGACAGCGTGTACTGCGGCGAACA 960
 Db 901 CCACCTACGACACCCGCGAGCGTTGCGCGTGCACGACAGCGTGTACTGCGGCGAACA 960
 QY 961 TGCCCGCTCGGTTGCGGAAAGCTGTGACCTACGCGCTGTGACCAAGCGGAGTGGCGTATG 1020
 Db 961 TGCCCGCTCGGTTGCGGAAAGCTGTGACCTACGCGCTGTGACCAAGCGGAGTGGCGTATG 1020
 QY 1021 TGCTGACCTTGGCGACCATGCTGGGCGGCGGCTGCGGCTGCAATTCGCGACTAGCA 1080
 Db 1021 TGCTGACCTTGGCGACCATGCTGGGCGGCGGCTGCGGCTGCAATTCGCGACTAGCA 1080
 QY 1081 AAGGCTTTTCACGACGAGAGGGCGTTACTGTCCGAAAGGGGTGGCCACGACCTGGGG 1140
 Db 1081 AAGGCTTTTCACGACGAGAGGGCGTTACTGTCCGAAAGGGGTGGCCACGACCTGGGG 1140
 QY 1141 TGCCGTTACCGAGAGCGCCGCGAGCGTGTGCTGACTCTCGGCGCTGTACCGCGAGC 1200
 Db 1141 TGCCGTTACCGAGAGCGCCGCGAGCGTGTGCTGACTCTCGGCGCTGTACCGCGAGC 1200
 QY 1201 ACAGCTCGGAGTAGGAGGAGGATGATG 1229
 Db 1201 ACAGCTCGGAGTAGGAGGAGGATGATG 1229

RESULT 6
 A89750 1229 bp DNA PAT 22-JAN-2000
 LOCUS A89750
 DEFINITION Sequence 7 from Patent WO9832862.
 ACCESSION A89750
 VERSION A89750.1 GI:6738284
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1229)
 AUTHORS Flohe, L. and Singh, M.
 TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
 JOURNAL Patent: WO 9832862-A 30-JUL-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source 1..1229
 Location/Qualifiers
 /organism="unidentified"
 /db_xref="taxon:32644"

BASE COUNT 236 a 391 c 382 g 219 t 1 others
 ORIGIN

Query Match 99.2%; Score 1225.4; DB 5; Length 1229;
 Best Local Similarity 99.8%; Pred. No. 1,1e-165;
 Matches 1229; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 ATCTTGAGATTAATGAACTTTCTTCACTGAAAGGTACAGATGAGAGGGGTATC 60
DB 1 ATCTTGAGATTAATGAACTTTCTTCACTGAAAGGTACAGATGAGAGGGGTATC 60
QY 61 ATGGGCGGTATTCGACGAGAGCAAAAACAAAGCAATCCGGTGGCCATCACCCG 120
DB 61 ATGGGCGGTATTCGACGAGAGCAAAAACAAAGCAATCCGGTGGCCATCACCCG 120
QY 121 GCGGCGGTGCGGGAAGTAAACCGTGTGGCCATGAGTGTCTACGAGGAGGTGCGGA 180
DB 121 GCGGCGGTGCGGGAAGTAAACCGTGTGGCCATGAGTGTCTACGAGGAGGTGCGGA 180
QY 181 GAGGCGGTGCGGTATTCACGAGAGCGGATTTCAAGGCGGAGCGCAACTGGTGGCACC 240
DB 181 GAGGCGGTGCGGTATTCACGAGAGCGGATTTCAAGGCGGAGCGCAACTGGTGGCACC 240
QY 241 GCCGACAGAGTGTGGGCGGACGCTGATTTATTTCTCAAGTCAAGAACGATAGCGGCG 300
DB 241 GCCGACAGAGTGTGGGCGGACGCTGATTTATTTCTCAAGTCAAGAACGATAGCGGCG 300
QY 301 GAATACGCGCGCGCTGCGAGAGCGGCGATCTGTTCAGCTT TTGATTTGGCGCGTAC 360
DB 301 GAATACGCGCGCGCTGCGAGAGCGGCGATCTGTTCAGCTTCTTGTGATTTGGCGCGTAC 360
QY 361 GTGCTTGACAGCATCGTGTGTGATTCGCGACACAGTCAATTGCTTACGAGACCGTCC 420
DB 361 GTGCTTGACAGCATCGTGTGTGATTCGCGACACAGTCAATTGCTTACGAGACCGTCC 420
QY 421 AGACGCGCGAGCGGCGACATACCCCTGCTTCCCGCATGAGCAATGCGCGGTGACACG 480
DB 421 AGACGCGCGAGCGGCGACATACCCCTGCTTCCCGCATGAGCAATGCGCGGTGACACG 480
QY 481 CGGCGAGGTGGCGCTTACACCTGATGCGAACCAAGGGGCGCGGTGTGCTGATGG 540
DB 481 CGGCGAGGTGGCGCTTACACCTGATGCGAACCAAGGGGCGCGGTGTGCTGATGG 540
QY 541 GCGGCGGTGCGGCGGTGCGAACCGGCGAGCTGTGATTCGCGCGCGGCGACCGCGGCT 600
DB 541 GCGGCGGTGCGGCGGTGCGAACCGGCGAGCTGTGATTCGCGCGCGGCGACCGCGGCT 600
QY 601 ACAAGCGACCGCGCATCGCGAACGGGCGGCGACGCTTACGCTTACAGATCAACA 660
DB 601 ACAAGCGACCGCGCATCGCGAACGGGCGGCGACGCTTACGCTTACAGATCAACA 660
QY 661 TCGACAAACTTGGGCAACTGAGCGGCGAGTGTGGCGCGA. ACACCTGCTACTCAT 720
DB 661 TCGACAAACTTGGGCAACTGAGCGGCGAGTGTGGCGCGA. ACACCTGCTACTCAT 720
QY 721 CGGCGTACGAGCTGAGGGGCGGCGTCAAAAGTGGCGGAGTGGGCGGCGTGG 780
DB 721 CGGCGTACGAGCTGAGGGGCGGCGTCAAAAGTGGCGGAGTGGGCGGCGTGG 780
QY 781 TCGCAGGCGCAAGGCAACCAAAATAGTCTGCAATTCACCTGTGGCGCATGAAACAG 840
DB 781 TCGCAGGCGCAAGGCAACCAAAATAGTCTGCAATTCACCTGTGGCGCATGAAACAG 840
QY 841 GTGCGGTATCTGTGATATAGCATGACAGGCGGCGTGTTCGAAAGGCTACAGCGGA 900
DB 841 GTGCGGTATCTGTGATATAGCATGACAGGCGGCGTGTTCGAAAGGCTACAGCGGA 900
QY 901 CCACCTACGACACCGAGCTTGGCGGTGACAGACAGCTGTTTCTGCTGTGGCGACA 960
DB 901 CCACCTACGACACCGAGCTTGGCGGTGACAGACAGCTGTTTCTGCTGTGGCGACA 960
QY 961 TGGCGCGTGTGGCGGAGAGCTGACCTACCGCTGACCAAGCGGAGATGCCGTATG 1020
DB 961 TGGCGCGTGTGGCGGAGAGCTGACCTACCGCTGACCAAGCGGAGATGCCGTATG 1020
QY 1021 TGGTGGAGCTTGGCGACCATGGGTGGGCGGCGGTGGCGGTGCAATTCGGGCACTAGCA 1080
  
```

```

DB 1021 TGGTGGAGCTTGGCGACCATGGGTGGGCGGCGGTGGCGGTGCAATTCGGGCACTAGCA 1080
QY 1081 AAGGCTTTTCAGACGACGAGAGGGGCGTACTGTCCGAACGGGTGGCCACGACCTGGGGG 1140
DB 1081 AAGGCTTTTCAGACGACGAGAGGGGCGTACTGTCCGAACGGGTGGCCACGACCTGGGGG 1140
QY 1141 TGGCGTTCACGAGCGCGCGGCGGCGTGTGGCGTGTGCTGCTGCGGCGCTGTACGGCCGANC 1200
DB 1141 TGGCGTTCACGAGCGCGCGGCGGCGTGTGGCGTGTGCTGCTGCGGCGCTGTACGGCCGAGC 1200
QY 1201 ACAGCTCGGAGTAAAGGAGGAGGATG 1229
DB 1201 ACAGCTCNGAGTAAAGGAGGAGGATG 1229
  
```

```

RESULT 7
A87613 1236 bp DNA PAT 22-JAN-2000
LOCUS A87613 Sequence 10 from Patent WO9836089.
DEFINITION A87613
ACCESSION A87613
VERSION A87613.1 GI:6736253
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 1236)
AUTHORS Flohe L. and Sinth M.
TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL Patent: WO 9836089-A 20-Aug-1998;
FLOHE LEOPOLD (DE); SINH MAHAVIR (DE)
FEATURES
source 1..1236
BASE COUNT 236 a 395 c 385 g 220 t
ORIGIN
  
```

Query Match 99.0%; Score 1223; DB 5; Length 1236;
 Best Local Similarity 99.8%; Pred. No. 2,5e-165;
 Matches 1234; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

QY 1 ATCTTGAGATTAATGAACTTTCTTCACTGAAAGGTACAGATGAGAGGGGTATC 60
DB 1 ATCTTGAGATTAATGAACTTTCTTCACTGAAAGGTACAGATGAGAGGGGTATC 60
QY 61 ATGGGCGGTATTCGACGAGAGCGGATTTCAAGGCGGAGCGCAATTCGGGCGGCGG 120
DB 61 ATGGGCGGTATTCGACGAGAGCGGATTTCAAGGCGGAGCGCAATTCGGGCGGCGG 120
QY 121 GCGGCGGTGCGGGAAGTAAACCGTGTGGCCATGAGTGTCTACAGGAGGTGCGGA 180
DB 121 GCGGCGGTGCGGGAAGTAAACCGTGTGGCCATGAGTGTCTACAGGAGGTGCGGA 180
QY 181 GAGGCGTGTGATACAGGAGCGGATTTGAAGGCGGAGGCGGCAACTGTGGCGACC 240
DB 181 GAGGCGTGTGATACAGGAGCGGATTTGAAGGCGGAGGCGGCAACTGTGGCGACC 240
QY 241 GCCGACAGAGTGTGGGCGAGCGAGTATTTGCTCAAGGTCAAGAACGAGATAGGCGG 300
DB 241 GCCGACAGAGTGTGGGCGAGCGAGTATTTGCTCAAGGTCAAGAACGAGATAGGCGG 300
QY 301 GAATACGCGCGCTGCGAGACGCGG. GATCTTGTTCACGCTTCTTGCATTTGGCGGCTCA 359
DB 301 GAATACGCGCGCTGCGAGACGCGG. GATCTTGTTCACGCTTCTTGCATTTGGCGGCTCA 360
QY 360 GGTGCTTACAGCATGCGTGTGTGATTCGCGGACGACGTAATTTGCTTACGAGACGCTC 419
DB 361 GGTGCTTACAGCATGCGTGTGTGATTCGCGGACGACGTAATTTGCTTACGAGACGCTC 420
QY 420 CAGACCGCGGAGCGGCGACTACCCCTGCTTCCCGATGAGCGAAGTGGCGGTGCACTC 479
  
```

Db 421 CAGACCGCGACGCGGCACTACCCCTGCTTGCCCGCATGAGCAAGTCCGGTGCATC 480
 QY 480 GCGGCCAGGTTGGCGCTTACCACTGATGCGCAACCAAGGGGGCGGGTGTGCTGATG 539
 Db 481 GCGGCCAGGTTGGCGCTTACCACTGATGCGCAACCAAGGGGGCGGGTGTGCTGATG 540
 QY 540 GCGGGGGTGGCGGCTGCAAGCGGCGAGAGTGTGTGATTCGGCGCGGCGGCGGCGG 599
 Db 541 GCGGGGGTGGCGGCTGCAAGCGGCGAGAGTGTGTGATTCGGCGCGGCGGCGGCGG 600
 QY 600 TACACGCAAGCGCGCATCGCCAAAGCGATGGGCGGCGGCGGCGGCGGCGGCGGCGG 659
 Db 601 TACACGCAAGCGCGCATCGCCAAAGCGATGGGCGGCGGCGGCGGCGGCGGCGGCGG 660
 QY 660 ATGACAAATTGGGCACTGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 719
 Db 661 ATGACAAATTGGGCACTGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
 QY 720 TCGGCTACGAGCTCGAGGCTGCGTCAACGTCGCGGCGGCGGCGGCGGCGGCGGCGG 779
 Db 721 TCGGCTACGAGCTCGAGGCTGCGTCAACGTCGCGGCGGCGGCGGCGGCGGCGGCGG 780
 QY 780 GTGCGAGGCGCGCAAGGCGCAATATGTCGAATTGCTGCGGCGGCGGCGGCGGCGGCGG 839
 Db 781 GTGCGAGGCGCGCAAGGCGCAATATGTCGAATTGCTGCGGCGGCGGCGGCGGCGGCGG 840
 QY 840 GTGCGGCTGCTGCTGATATAGCATTCGACAGGCGGCGGCGGCGGCGGCGGCGGCGG 899
 Db 841 GTGCGGCTGCTGCTGATATAGCATTCGACAGGCGGCGGCGGCGGCGGCGGCGGCGG 900
 QY 900 ACCAGCTACAGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 959
 Db 901 ACCAGCTACAGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
 QY 960 ATGCGCGGCTGCTGCTGATATAGCATTCGACAGGCGGCGGCGGCGGCGGCGGCGGCGG 1019
 Db 961 ATGCGCGGCTGCTGCTGATATAGCATTCGACAGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
 QY 1020 GTGCTGAGGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1079
 Db 1021 GTGCTGAGGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
 QY 1080 AAAGGCTTTCGAGCAGCAAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1139
 Db 1081 AAAGGCTTTCGAGCAGCAAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
 QY 1140 GTGCGGCTTACGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1199
 Db 1141 GTGCGGCTTACGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
 QY 1200 CACACGTCGGGAGTAAGGGAAGGATGATTCGGCC 1235
 Db 1201 CACACGTCGGGAGTAAGGGAAGGATGATTCGGCC 1236

RESULT 8
 A89753
 LOCUS A89753 1236 bp DNA
 DEFINITION Sequence 10 from Patient MO9832862.
 ACCESSION A89753
 VERSION A89753.1 GI:6738287
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1236)
 AUTHORS Flohe, L. and Singh, M.
 TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
 JOURNAL Patent: WO 9832862-A 30-JUL-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source 1..1236
 /organism="unidentified"

BASE COUNT 236 a 395 c 385 g 220 t
 ORIGIN
 Query Match 99.0%; Score 1223; DB 5; Length 1236;
 Best local similarity 99.8%; Pred. No. 2,5e-165;
 Matches 1234; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATCTTCAGATTAAATGAACCTTCTTACACTGAAAGCGTACAGTATCGAAGGGGTAATC 60
 Db 1 ATCTTCAGATTAAATGAACCTTCTTACACTGAAAGCGTACAGTATCGAAGGGGTAATC 60
 QY 61 ATGCGGCTGGGATATTCGACAGGACCAAAACAGAAATTCGGGGTGGCGATCACCCG 120
 Db 61 ATGCGGCTGGGATATTCGACAGGACCAAAACAGAAATTCGGGGTGGCGATCACCCG 120
 QY 121 GCGGCGTGGGAACTAACCGCTGCTGCGCATGAGTGTCTATCCAGGAGAGTGGCGGA 180
 Db 121 GCGGCGTGGGAACTAACCGCTGCTGCGCATGAGTGTCTATCCAGGAGAGTGGCGGA 180
 QY 181 GAGGCTGCGGCTATACGAGCGGCGGATTCAGGCGGCGGCGGCGGCGGCGGCGGCGG 240
 Db 181 GAGGCTGCGGCTATACGAGCGGCGGATTCAGGCGGCGGCGGCGGCGGCGGCGGCGG 240
 QY 241 GCGGACAGGTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
 Db 241 GCGGACAGGTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
 QY 301 GAATACGGCGCGCTGCGGACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 359
 Db 301 GAATACGGCGCGCTGCGGACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
 QY 360 CGTGTTCGACCGGATCGTTGTTGATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 419
 Db 361 CGTGTTCGACCGGATCGTTGTTGATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
 QY 420 CAGACCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 479
 Db 421 CAGACCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
 QY 480 GCGGCGGAGGTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 539
 Db 481 GCGGCGGAGGTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
 QY 540 GCGGCGGAGGTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 599
 Db 541 GCGGCGGAGGTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
 QY 600 TACACGCAAGCGCGCATTCGCGCAAGCGGATGGGCGGCGGCGGCGGCGGCGGCGGCGG 659
 Db 601 TACACGCAAGCGCGCATTCGCGCAAGCGGATGGGCGGCGGCGGCGGCGGCGGCGGCGG 660
 QY 660 ATGACAAATTGGGCACTGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 719
 Db 661 ATGACAAATTGGGCACTGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
 QY 720 TCGGCTTACGAGCTCGAGGCTGCGTCAACGTCGCGGCGGCGGCGGCGGCGGCGGCGG 779
 Db 721 TCGGCTTACGAGCTCGAGGCTGCGTCAACGTCGCGGCGGCGGCGGCGGCGGCGGCGG 780
 QY 780 GTGCGAGGCGCGCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 839
 Db 781 GTGCGAGGCGCGCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
 QY 840 GTGCGGCTGCTGCTGATATAGCATTCGACAGGCGGCGGCGGCGGCGGCGGCGGCGG 899
 Db 841 GTGCGGCTGCTGCTGATATAGCATTCGACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
 QY 900 ACCAGCTACAGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 959
 Db 901 ACCAGCTACAGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960

QY	960	ATGCGCGCTCGGGGCCGAAGACGTGACCTACGGCCTACCAACGCGACGATGCCGAT	101.9
QY	961	ATGCGCGCTCGGGGCCGAAGACGTGACCTACGGCCTACCAACGCGACGATGCCGAT	1020
QY	1020	GTGCTCGAGCTTGCAGCCATATGCTGCGGGGCGGTGCGGTCGAATCCGGCAGCTAACCC	107.9
Db	1021	GTGCTCGAGCTTGCAGCCATATGCTGCGGGGCGGTGCGGTCGAATCCGGCAGCTAACCC	1080
QY	1080	AAAGCTCTTTCGACGCGACGAAAGGGGGGTTACTCTCCGAACGGGTGGCCACGACCTGGGG	113.9
Db	1081	AAAGCTCTTTCGACGCGACGAAAGGGGGGTTACTCTCCGAACGGGTGGCCACGACCTGGGG	1140
QY	1140	GTGCGGTCACCGAGCCCGCCACAGCTGTGAGCTCTGCTGCGGCGCTGTTACCGCGAN	119.9
Db	1141	GTGCGGTCACCGAGCCCGCCACAGCTGTGAGCTCTGCTGCGGCGCTGTTACCGCGAG	1200
QY	1200	CACACGTCGGAGTAAGGGAAGGAGATGATGCGGCC	123.5
Db	1201	CACACGTCGGAGTAAGGGAAGGAGATGATGCGGCC	1236

RESULT	9		PAT	22-JAN-2000
A87607				
LOCUS	A87607	1237 bp	DNA	
DEFINITION	Sequence 4 from Patent WO9836089.			
ACCESSION	A87607			
VERSION	A87607.1	GI:5736247		
KEYWORDS	.			
SOURCE	unidentified.			
ORGANISM	unidentified.			

REFERENCE 1 (bases 1 to 1237)
AUTHORS Flohe, L. and Singh, M.
TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL Patent: WO 9836089-A 20-AUG-1998:

```

FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
location/Qualifiers
source
1. .1237

```

ORIGIN	BASE COUNT				
	236	a	386	g	221
			394	c	
					221
					t

Query Match	98.98;	Score 1221.4;	DB 5;	Length 1237;
Best Local Similarity	99.88;	Pred. No. 4.2e-165;		
Matches 1233;	Conservative	0;	Mismatches 2;	Indels 1;
				Gaps 1;

QY	1	ATCTGCAGATTAAATCGAACTTCTTCACTGAAGCGTACGTATCGAAGGGGTATC	60
Db	1	ATCTGCAGATTAAATCGAACTTCTTCACTGAAGCGTACGTATCGAAGGGGTATC	60
QY	61	ATGGCGCGCGGTATTCGACCGAGCAACCAAAACAACAATTCGGGGGTGCATCACCCG	120
Db	61	ATGGCGCGCGGTATTCGACCGAGCAACCAAAACAACAATTCGGGGGTGCATCACCCG	120
QY	121	GCCGGCGTGGGGAATTAACCGCTGTGGCCATGAGTGTCTATCCAGGAGGTCCGGA	180
Db	121	GCCGGCGTGGGGAATTAACCGCTGTGGCCATGAGTGTCTATCCAGGAGGTCCGGA	180
QY	181	GAGGCGTCTGCTATCACCGACGCGGGATTTCAAGCGGCGAGCGCGCAACTGTGCGCAC	240
Db	181	GAGGCGTCTGCTATCACCGACGCGGGATTTCAAGCGGCGAGCGCGCAACTGTGCGCAC	240
QY	241	GCCGACCAAGGTGTGGCGCGAAGCGATTTATGTCTCAAGGCAAGAACCGATATGCGCG	300
Db	241	GCCGACCAAGGTGTGGCGCGAAGCGATTTATGTCTCAAGGCAAGAACCGATATGCGCG	300
QY	301	GAATATGCGCGCGCTGCGACAGAGGCG - GATCTTGTTCAGGTCTTGCATTGGCGCGCTCA	359
Db	301	GAATATGCGCGCGCTGCGACAGAGGCGAGATCTTGTTCAGGTCTTGCATTGGCGCGCTCA	360
QY	360	CGTGGCTGCACCGATGCGGTGTGTGGATTCGCGGACACAGCTGATATTTGCCCTACGAGACGCTC	419

Db	361	CGTCTTGCACCCGATGCGTTGTGTGATTCGGACACGCTCAATTGGCTTAGAGACGGTC	420
QY	420	CAGACGCCGACGACGGCGCATACCCCTGCTTGGCCCGATGAGCGAAGTCGCCGGTGCATC	479
Db	421	CAGACGCCGACGACGGCGCATACCCCTGCTTGGCCCGATGAGCGAAGTCGCCGGTGCATC	480
QY	480	CGCCGCCAGGTTGGCGTTACCACTGATGCGCAACCAAGGGGGCCGGTGTGCTATG	539
Db	481	GGCCGCCAGGTTGGCGTTACCACTGATGCGCAACCAAGGGGGCCGGTGTGCTATG	540
QY	540	GGCGGGGTGCGCGCGCGTCGAAACCGGGCGAGCGTCGTATGATGGCGCGGACCGCGGC	599
Db	541	GGCGGGGTGCGCGCGCGTCGAAACCGGGCGAGCGTCGTATGATGGCGCGGACCGCGGC	600
QY	600	TACAAGCGACCCGCGATCGCCAAAGGCGATGGCGGACCGATTACGGTTCTAGACATCAAC	659
Db	601	TACAACCGACCCGCGATCGCCAAAGGCGATGGCGGACCGATTACGGTTCTAGACATCAAC	660
QY	660	ATGCACAACTTGGCGCACTCGACGCCCGAGTTCTGCGCGCGGATCAACACTCGTACTCA	719
Db	661	ATGCACAACTTGGCGCACTCGACGCCCGAGTTCTGCGCGCGGATCAACACTCGTACTCA	720
QY	720	TCGGCTACGAGAGTCGAGGGTGGCGTCAACAGTCGCGACCTGGTATTGGGGCGCTCTG	779
Db	721	TCGGCTACGAGAGTCGAGGGTGGCGTCAACAGTCGCGACCTGGTATTGGGGCGCTCTG	780
QY	780	GTCGCAGGCGCCAAAGCACCACAAATTAGTCTGAAATTCATTGTGCGGCATATGAAACA	839
Db	781	GTCGCAGGCGCCAAAGCACCACAAATTAGTCTGAAATTCATTGTGCGGCATATGAAACA	840
QY	840	GGTCGGTACTGGGGGATATAGCCATTCGACGACGAGGGGGGCTGTTGGAAGTCACGACCG	899
Db	841	GGTCGC TACTGGGATATAGCCATTCGACGACGAGGGGGGCTGTTGGAAGTCACGACCG	900
QY	900	ACCACTACGACGACCGGACGTTGGCGTGCAGACGACGCTGTTTACTGGTGGCGAAC	959
Db	901	ACCACTACGACGACCGGACGTTGGCGTGCAGACGACGCTGTTTACTGGTGGCGAAC	960
QY	960	ATGCGCGCTTCGGTGCAGAAAGCTGACCTACGCGCTGACCAACGAGAGATGCCGAT	1019
Db	961	ATGCGCGCTTCGGTGCAGAAAGCTGACCTACGCGCTGACCAACGAGAGATGCCGAT	1020
QY	1020	GTCGTCAGGCTGGCCACCACTGCTGGCGGGCGGCTGCGGCTGCAATCCGGCACTAGCC	1079
Db	1021	GTCGTCAGGCTGGCCACCACTGCTGGCGGGCGGCTGCGGCTGCAATCCGGCACTAGCC	1080
QY	1080	AAAGGCTTTTCGACGACGACGAGGGCGCTTACTGTCGAAAGGGGTGGCCACGACCTAGGG	1139
Db	1081	AAAGGCTTTTCGACGACGACGAGGGCGCTTACTGTCGAAAGGGGTGGCCACGACCTAGGG	1140
QY	1140	GTCGCGTTACCGAGCCCGCGACGCTGGCTGACTTCGCGCGCTGTTACGCGGAN	1199
Db	1141	GTCGCGTTACCGAGCCCGCGACGCTGGCTGACTTCGCGCGCTGTTACGCGGAG	1200
QY	1200	CACACGTCGGGAGTAAGGGAGGAGATGATGGGCC	1235
Db	1201	CACACGTCGGGAGTAAGGGAGGAGATGATGGGCC	1236

RESULT	10				
A89747					
LOCUS	A89747	1237 bp	DNA	PAT	22-JAN-2000
DEFINITION	Sequence 4 from Patent WO9832862.				
ACCESSION	A89747				
VERSION	A89747.1	GI:6738281			
KEYWORDS	.				
SOURCE	unidentified.				
ORGANISM	unidentified				
	unclassified.				
REFERENCE	1 (bases 1 to 1237)				
AUTHORS	Flohe,L. and Singh,N.				
TITLE	L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM				

CAUTION: In some cases we may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtc, or ttg) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

FEATURES

```

source
  1.56414
  /organism="Mycobacterium tuberculosis"
  /strain="H37Rv"
  /db_xref="taxon:1773"
  <1..>56414
  /note="fragment designated v002. Does not represent a physical clone"
  <1..233
  /organism="Mycobacterium tuberculosis"
  /strain="H37Rv"
  /db_xref="taxon:1773"
  /clone="Y154"
  /complement(3..527)
  /gene="recX"
  /complement(3..527)
  /gene="recX"
  /note="RV2738c, (MTV002.01c), len: 174 aa. recX, similar to eg. recX_PSEAE P37860 regulatory protein recX from Pseudomonas aeruginosa (153 aa), fasta scores: opt: 161 z-score: 257.2 E(): 3.6e-07, 30.7% identity in 137 aa overlap. Overlaps and extends CDS from overlapping cosmid MCY154.16c"
  /codon_start=-1
  /transl_table=11
  /product="recX"
  /protein_id="CA15532.1"
  /db_xref="GI:2624257"
  /db_xref="SPTREMBL:O33280"
  /translation="MTVSCPSPSTSESEBOARLCLRLTARSRTAEALAGLAKRGY PEDIGRVDRLAAGLVDDTDFADQWQWRAAKSKRALAEHLHAGVDDVITTVIGGIDAGARGRAKELVRLRLREVLIIDGTGDEANVSRLVAMLRGCGTACEV VIELAERERRRNV"
  /complement(493..2865)
  /gene="recA"
  /complement(493..2865)
  /gene="recA"
  /note="RV2737c, (MTV002.02c), len: 790 aa. recA, identical to RECA_MCTU P26345 recA protein (790 aa). Contains self-splicing protein element (intron) from 2294 to 969 (c), similar to intron II from TR:E332317 (EMBL:Y13030) DNA-directed DNA polymerase (EC 2.7.7.7) from Thermococcus sp. (1829 aa), fasta scores: opt: 81 z-score: 235.2 E(): 6e-06, 24.6% identity in 183 aa overlap. Contains PS00017 ATP/GTP-binding site motif A (P-loop), PS00321 recA signature, and PS00881 Protein splicing signature. See Davis et al. (1992) Cell 11(2):201-210"
  /codon_start=-1
  /transl_table=11
  /product="recA"
  /protein_id="CA15533.1"
  /db_xref="GI:2624258"
  /db_xref="SWISS-PROT:P26345"
  /translation="MTQTPDERKALELAVNQIEKSYGKGSVMRLGDEARQISVIFPG SIADVALIGIGLGRGVIEIGPSSSGI TVALHAYANAAGVVAAPIDADHADP DYAKKIDVDSLSOPDTGEALLETAL ALISGALDIIVDSVALVRALESEM GDSHVIGQARLMGALRMGALNGSTGTAIFINLRDKIGVMEGSEPTTGKALKE YASRMVRRVETLKDGTNAVGNTRPVKKNCLAGSTIIPDPVGTTHRIDGAG RKLHVVAAKAGDTLHARPVVSNPDGTQVYIGIRAGCAIVATPDKHKTLEYGARA AGELRKGDRAVQPRFDGDSAPIPADHARLGLYLIGDRDVGWGGKIPININQV ALIDVYRIATATGCAHPQGRISLAHRRGERNGVADLCQAGIYGLKAMEKTIEN WFEPPDIAADIVGNLIFGLFESDGVWSREDTGALRVGTTTSEQLQHMLLRGCV GSTVYRDPQKRSIVNGRIISGRKQVFEVRIAGDMNTAAFEASVPMQPRAAALLO AIPATGRRGSGATYLAEMTDVAVNYEDRGVTNOEAAAMVAGVSGPRGGMKVO LGASRLRRDRAVALADLDKDFLHMDLAEEELRSVIREVYPTRRARPDLEVEELHTL VAEGVVHNCSPPKQAEFDLILGKGISRESGLIDKGVDDGLIRKSGAMTTEYGEOLG OGKGNANFLVENADVADEIEKIKELIGAVVTDPSNDGVLPAVDE"

```

```

misc_feature
  complement(790..2112)
  /gene="recA"
  /note="recA intein"
  complement(790..807)
  /gene="recA"
  /note="PS00881 protein splicing signature"
  complement(2197..2223)
  /gene="recA"
  /note="PS00321 recA signature"
  complement(2644..2667)
  /gene="recA"
  /note="PS00017 ATP/GTP-binding site motif A (P-loop)"
  complement(2873..2877)
  /note="possible RBS upstream of RV2737c"
  complement(3247..3453)
  /gene="RV2738c"
  complement(3247..3453)
  /gene="RV2738c"
  /note="RV2738c, (MTV002.03c), len: 68 aa. Unknown, N-terminus is highly similar to the N-terminus of the upstream ORF MTV002.07c (78.4% identity in 37 aa overlap); also similar to AL020958|SC4H8.5 Streptomyces coelicolor cosmid 4H8 (64aa) opt: 185 z-score: 283.5 E(): 2.9e-08; 39.7% identity in 63 aa overlap"
  /codon_start=-1
  /transl_table=11
  /product="hypothetical protein RV2738c"
  /protein_id="CA15534.1"
  /db_xref="GI:2624260"
  /db_xref="SPTREMBL:O33281"
  /translation="MLAGVRLTEFEHERALHFGAAYGSSVLDHVLVGFQDSRAQAAL EDGVPRDVRALCADFVPRDRW"
  complement(3464..4530)
  /gene="RV2739c"
  complement(3464..4530)
  /gene="RV2739c"
  /note="RV2739c, (MTV002.04c), len: 388 aa. Probable transfe use, similar to eg. TRC 1560 (EMBL:L28170) Pseudomonas aeruginosa thiamosyl transferase (426 aa) fasta scores: opt: 178 z-score: 226.3 E(): 1.9e-05, 25.9% identity in 425 aa overlap. Equivalent to Mycobacterium leprae protein MCB33.02c (392 aa), fasta scores 9P1294723|MCB33.2 Mycobacterium leprae cosmid B33 opt: 2112 z-score: 2364.5 E(): 0; 80.9% identity in 388 aa overlap"
  /codon_start=-1
  /transl_table=11
  /product="hypothetical protein RV2739c"
  /protein_id="CA15535.1"
  /db_xref="GI:2624261"
  /db_xref="SPTREMBL:O33282"
  /translation="KRVAVAGPDDGHSFPAIALCQRRRAADPTLTGVEYLEAR AAGIDAVEIDGLATDRDIDADARLHRRRAQAVANVRLALEPELVSQVITACGG MAAEILIPWELNHPHLYLPSKGLPPIGSGLAAGTGRRLRATKALGRSMRAG LRQRAVREVEILPARDPGLRLIATIPALEVPPDPAEAVVVGPLHFPDTRVLA IPAGTGPVVAVPSTALGTAGLTVALQSITPGETVSGSRLVYSLGADLVTPPM AVAGIGSQAELLIRADIVICGGHGMVAKTTLIAGVPMVVVGSGDQMEIARVYRQGS AVILRPTNDALVAVNVEYLSRPREARRRAAASVAAADPVRCHALALAG"
  4674..5123
  /gene="RV2740"
  4674..5123
  /gene="RV2740"
  /note="RV2740, (MTV002.05), len: 149 aa. unknown, similar to eg. leprae hypothetical protein TR:049850 (EMBL:294723) MCB33.03 (82235_F3_140) (178 aa), fasta scores: opt: 498 z-score: 538.9 E(): 7.3e-23, 51.6% identity in 161 aa overlap"
  /codon_start=-1
  /transl_table=11
  /product="hypothetical protein RV2740"
  /protein_id="CA15536.1"
  /db_xref="GI:2624262"
  /db_xref="SPTREMBL:O33283"

```


Db 781 GTGCCAGGCGCCAGGAGCCCAATTAGTCTGAAATTCATTCTGCGCATATGAAACA 840
QY 840 GTGCGGACTAGTGTATATAGCATCGACAGGCGGCTGTTTGAAGGCTCAGACGCG 899
Db 841 GTGCGGACTAGTGTATATAGCATCGACAGGCGGCTGTTTGAAGGCTCAGACGCG 900
QY 900 ACCACCTACGACCAACCCGACGCTGCGGTGACAGACAGCTGTTTACTGCGTGGCGAAC 959
Db 901 ACCACCTACGACCAACCCGACGCTGCGGTGACAGACAGCTGTTTACTGCGTGGCGAAC 960
QY 960 ATGCGCGCGGCTGCGGCAAGACGCTGCGGTGACAGACAGCTGTTTACTGCGTGGCGAAC 1019
Db 961 ATGCGCGCGGCTGCGGCAAGACGCTGCGGTGACAGACAGCTGTTTACTGCGTGGCGAAC 1020
QY 1020 GTGCTGAGCTTGGCGACCATGCTGCGGCGGCGGCTGCGGTGACATCGGACATACC 1079
Db 1021 GTGCTGAGCTTGGCGACCATGCTGCGGCGGCGGCTGCGGTGACATCGGACATACC 1080
QY 1080 AAAGGCTTTTGGACGACGAAAGGCGGCTTACTGTCGAAAGGCGTGGCCACCGACCTGGGG 1139
Db 1081 AAAGGCTTTTGGACGACGAAAGGCGGCTTACTGTCGAAAGGCGTGGCCACCGACCTGGGG 1140
QY 1140 GTGCGGCTACGAGCGCGCGGCGGCGGCTGCTGCGGTGACCTGCGGCGGCTGTTAGCGCGAN 1199
Db 1141 GTGCGGCTACGAGCGCGCGGCGGCGGCTGCTGCGGTGACCTGCGGCGGCTGTTAGCGCGAG 1200
QY 1200 CACACGTGCGGAGTAAGGAGAGCATGATGTCGCG 1234
Db 1201 CACACGTGCGGAGTAAGGAGAGCATGATGTCGCG 1235

RESULT 14
LOCUS A87608 1228 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 5 from Patent M09836089.
ACCESSION A87608.1 GI:6736248
VERSION A87608.1 GI:6736248
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1228)
AUTHORS Flohe, L. and Singh, M.
TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL Patent: WO 9836089-A-20-AUG-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
source 1..1228
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 236 a 391 c 382 g 219 t
ORIGIN

Query Match 98.3%; Score 1213.4; DB 5; Length 1228;
Best Local Similarity 99.8%; Pired. No. 5.8e-164;
Matches 1225; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATCTTGAGATTATCGAAGCTTCTTCACTGAAGCTAAGATCGAAGGGGTATC 60
Db 1 ATCTTGAGATTATCGAAGCTTCTTCACTGAAGCTAAGATCGAAGGGGTATC 60
QY 61 ATGCGCGGCTGCTGCGGACGACGAAACAAAGAAATTCGCGGCTGCGGCTATCACC 120
Db 61 ATGCGCGGCTGCTGCGGACGACGAAACAAAGAAATTCGCGGCTGCGGCTATCACC 120
QY 121 GCGCGGCTGCGGAGCTAACCCTGCTGCGGCTATGAGTGTCTATCAGGCGAGTGGCGGA 180
Db 121 GCGCGGCTGCGGAGCTAACCCTGCTGCGGCTATGAGTGTCTATCAGGCGAGTGGCGGA 180
QY 181 GAGGCGTGGCTATCAGCGAGCGGATTCAGAGGCGGCGGCGGCAACGTGGCGCAC 240
Db 181 GAGGCGTGGCTATCAGCGAGCGGATTCAGAGGCGGCGGCGGCAACGTGGCGCAC 240

QY 241 GCCGACAGGCTGCGGCGGACGCTGATTTATGCTCAAGGCAAGAACCGATAGCGCG 300
Db 241 GCCGACAGGCTGCGGCGGACGCTGATTTATGCTCAAGGCAAGAACCGATAGCGCG 300
QY 301 GAATACGCGCGCTGCGGACAGCGGC -GATCTTGTTCAGCTTGTTCATTTGGCGGCTCA 359
Db 301 GAATACGCGCGCTGCGGACAGCGGCATCTTGTTCAGCTTGTTCATTTGGCGGCTCA 360
QY 360 GGTCTTGCACCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 419
Db 361 GGTCTTGCACCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 420 CAGACCGCGGCGGACGCTACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 479
Db 421 CAGACCGCGGCGGACGCTACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 480 GCGCGGAGGCTGCGGCTTACCACTGATGCGAACCAGGCGGCGGCTGCTGCTGCTGCTGCT 539
Db 481 GCGCGGAGGCTGCGGCTTACCACTGATGCGAACCAGGCGGCGGCTGCTGCTGCTGCTGCT 540
QY 540 GCGCGGAGGCTGCGGCTTACCACTGATGCGAACCAGGCGGCGGCTGCTGCTGCTGCTGCT 599
Db 541 GCGCGGAGGCTGCGGCTTACCACTGATGCGAACCAGGCGGCGGCTGCTGCTGCTGCTGCT 600
QY 600 TACACGAGCGGCGGCTACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 659
Db 601 TACACGAGCGGCGGCTACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 660
QY 660 ATGACAACTTGGCACTGACGCGGAGTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 719
Db 661 ATGACAACTTGGCACTGACGCGGAGTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
QY 720 TCGGC TACGAGCTGAGAGGTCGCGTAAACGTCGCGGCGGCGGCGGCGGCGGCGGCGG 779
Db 721 TCGGC TACGAGCTGAGAGGTCGCGTAAACGTCGCGGCGGCGGCGGCGGCGGCGGCGG 780
QY 780 GTGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 839
Db 781 GTGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
QY 840 GTGCGGAGCTGATGATATAGCATGACGACGAGGCGGCTGTTTGAAGGCTCAGACGCG 899
Db 841 GTGCGGAGCTGATGATATAGCATGACGACGAGGCGGCTGTTTGAAGGCTCAGACGCG 900
QY 900 ACCACCTACGACCAACCCGACGCTGCGGTGACAGACAGCTGTTTACTGCGTGGCGAAC 959
Db 901 ACCACCTACGACCAACCCGACGCTGCGGTGACAGACAGCTGTTTACTGCGTGGCGAAC 960
QY 960 ATGCGCGGCTGCGGCGGAGAGCTGACCTAGCGGCTGACCAACGAGAGATGCGGTAT 1019
Db 961 ATGCGCGGCTGCGGCGGAGAGCTGACCTAGCGGCTGACCAACGAGAGATGCGGTAT 1020
QY 1020 GTGCTGAGCTTGGCGACCATGCTGCGGCGGCGGCGGCTGCGGTGACATCGGACATACC 1079
Db 1021 GTGCTGAGCTTGGCGACCATGCTGCGGCGGCGGCGGCTGCGGTGACATCGGACATACC 1080
QY 1080 AAAGGCTTTTGGACGACGAAAGGCGGCTTACTGTCGAAAGGCGTGGCCACCGACCTGGGG 1139
Db 1081 AAAGGCTTTTGGACGACGAAAGGCGGCTTACTGTCGAAAGGCGTGGCCACCGACCTGGGG 1140
QY 1140 GTGCGGCTACGAGCGCGCGGCGGCGGCTGCTGCGGTGACCTGCGGCGGCTGTTAGCGCGAN 1199
Db 1141 GTGCGGCTACGAGCGCGCGGCGGCGGCTGCTGCGGTGACCTGCGGCGGCTGTTAGCGCGAG 1200
QY 1200 CACACGTGCGGAGTAAGGAGAGCATG 1228
Db 1201 CACACGTGCGGAGTAAGGAGAGCATG 1228

RESULT 15
LOCUS A89748 1228 bp DNA PAT 22-JAN-2000

DEFINITION Sequence 5 from Patent WO9832862.
 ACCESSION A89748
 VERSION A89748.1 GI:6738282
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1228)
 FLOHE, L. and SINGH, M.
 TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
 JOURNAL Patent: WO 9832862-A 30-JUL-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source 1..1228
 Location/Qualifiers
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 236 a 391 c 382 g 219 t
 ORIGIN

Query Match 98.3%; Score 1213.4; DB 5; Length 1228;
 Best Local Similarity 99.8%; Pred. No. 5.8e-164;
 Matches 1225; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATCTTGAGATTATCGAAGCTTCTTCTACACTGAAGCGTACAGTATCGAGAGGGTATATC 60
 DB 1 ATCTTGAGATTATCGAAGCTTCTTCTACACTGAAGCGTACAGTATCGAGAGGGTATATC 60
 QY 61 ATGGGCGTGGTATTCGACCGAGACCAAAACAGAAATTCGGGGTGGCATACCCCG 120
 DB 61 ATGGGCGTGGTATTCGACCGAGACCAAAACAGAAATTCGGGGTGGCATACCCCG 120
 QY 121 GCCGCGTCGCGGAACTACCCGTCGTGGCCATGAGGTGCTCATCCAGGAGGTGCCGA 180
 DB 121 GCCGCGTCGCGGAACTACCCGTCGTGGCCATGAGGTGCTCATCCAGGAGGTGCCGA 180
 QY 181 GAGGCTCGGCTATCACCGAGCGGATTTCAAGCGCGGCGCAACTGGTGGCAACC 240
 DB 181 GAGGCTCGGCTATCACCGAGCGGATTTCAAGCGCGGCGCAACTGGTGGCAACC 240
 QY 241 GCGGACGAGGTGGGCGGAGCGGTATTTATGCTCAAGGTCAAGAAGCGATAGCGCG 300
 DB 241 GCGGACGAGGTGGGCGGAGCGGTATTTATGCTCAAGGTCAAGAAGCGATAGCGCG 300
 QY 301 GAATACGCGCGCTGCGACACAGGCG -GATCTTGTTCACGTTCTTGCATTGGCCGCTCA 359
 DB 301 GAATACGCGCGCTGCGACACAGGCGAGATCTTGTTCACGTTCTTGCATTGGCCGCTCA 359
 QY 360 CGTCTTGACCGCATGGGTGTGATTCGGGCAACGATGCTCAATGGCTACGAGACCGTC 419
 DB 360 CGTCTTGACCGCATGGGTGTGATTCGGGCAACGATGCTCAATGGCTACGAGACCGTC 419
 QY 420 CAGACCGCGAGCGGAGCTACCTCTTGGCCGATGAGCGAAGTCCGCGTGAATC 479
 DB 420 CAGACCGCGAGCGGAGCTACCTCTTGGCCGATGAGCGAAGTCCGCGTGAATC 479
 QY 480 GCCGCGCAGGTTGGCGTTACCACTGATGCGAACCAAGGGGCGCGGCTGTGATG 539
 DB 480 GCCGCGCAGGTTGGCGTTACCACTGATGCGAACCAAGGGGCGCGGCTGTGATG 539
 QY 540 GCGCGCGGTTGGCGGCTGCGAACCGCGGAGCTGCTGATGCGCGCGCGCGCGC 599
 DB 540 GCGCGCGGTTGGCGGCTGCGAACCGCGGAGCTGCTGATGCGCGCGCGCGCGC 599
 QY 541 GCGCGGTTGGCGGCTGCGAACCGCGGAGCTGCTGATGCGCGCGCGCGCGC 600
 DB 541 GCGCGGTTGGCGGCTGCGAACCGCGGAGCTGCTGATGCGCGCGCGCGCGC 600
 QY 600 TACAAGCAGCGCGCATGCGCAAGCGCATGGGCGCGACCGTTACGTTCTAGACATCAAC 659
 DB 600 TACAAGCAGCGCGCATGCGCAAGCGCATGGGCGCGACCGTTACGTTCTAGACATCAAC 659
 QY 660 ATGACAAACTTGGGCAACTGACGCGGAGTTGGGCGCGGATCCACACTCGCTACTCA 719
 DB 660 ATGACAAACTTGGGCAACTGACGCGGAGTTGGGCGCGGATCCACACTCGCTACTCA 719
 QY 720 ATGACAAACTTGGGCAACTGACGCGGAGTTGGGCGCGGATCCACACTCGCTACTCA 720
 DB 720 ATGACAAACTTGGGCAACTGACGCGGAGTTGGGCGCGGATCCACACTCGCTACTCA 720
 QY 720 TCGGCTACGAGCTCGAGGTTGGCTCAAGCTGCGGAGCTGCTGATGGGCGCTCTG 779
 DB 720 TCGGCTACGAGCTCGAGGTTGGCTCAAGCTGCGGAGCTGCTGATGGGCGCTCTG 779

DB 721 TCGGCTACGAGCTCGAGGTTGGCTCAAGCTGCGGAGCTGCTGATGGGCGCTCTG 780
 QY 780 GTGCCAGGCGCCAGGACCAAAATTAAGTCTGATTTACTGTTGCGCATATGAACCA 839
 DB 781 GTGCCAGGCGCCAGGACCAAAATTAAGTCTGATTTACTGTTGCGCATATGAACCA 840
 QY 840 GTGCCAGTCTGATTTAGCATGACGAGGCGGCTGTTGGAAGGCTACGACCG 899
 DB 841 GTGCCAGTCTGATTTAGCATGACGAGGCGGCTGTTGGAAGGCTACGACCG 900
 QY 900 ACCACCTACGACCCAGCGTTCCGCGTGCACGACAGCGTGTCTTACTGCGTGCAGAC 959
 DB 901 ACCACCTACGACCCAGCGTTCCGCGTGCACGACAGCGTGTCTTACTGCGTGCAGAC 960
 QY 960 ATGCCCGCTCGTGGCGGAGAGCGTACGACTACGCGGCTGACCAAGCGGATGCGCTAT 1019
 DB 961 ATGCCCGCTCGTGGCGGAGAGCGTACGACTACGCGGCTGACCAAGCGGATGCGCTAT 1020
 QY 1020 GTGCTGAGCTTGGCGACCATGAGTGGCGGCGGCTGCGGCTGATTCGATCCGACTAGCC 1079
 DB 1021 GTGCTGAGCTTGGCGACCATGAGTGGCGGCGGCTGCGGCTGATTCGATCCGACTAGCC 1080
 QY 1080 AAAGTCTTTGACGACGACGAGGCGGCTTACTGTCCGAACGGGTGGCCACCACTGGGG 1139
 DB 1081 AAAGTCTTTGACGACGACGAGGCGGCTTACTGTCCGAACGGGTGGCCACCACTGGGG 1140
 QY 1140 GTGCGTTACGAGCGCGCCAGCGTGTGCGCTGACTCTGCGCGCTGTTACGCGG 1199
 DB 1141 GTGCGTTACGAGCGCGCCAGCGTGTGCGCTGACTCTGCGCGCTGTTACGCGG 1200
 QY 1200 CACACGTCGAGGAGTAAGGAGCGATGA 1227
 DB 1201 CACACGTCGAGGAGTAAGGAGCGATGA 1228

Search completed: June 22, 2000, 14:54:29
 Job time: 17684 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 15:08:21 ; Search time 458.59 Seconds
(without alignments)
673.777 Million cell updates/sec

Title: US-09-362-485-6

Perfect score: 1235
Sequence: 1 ATCTTGCGATTAATCGAAC.....GGCAGCGATGATGCGGCC 1235

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 31185 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1207	97.7	1245	1 V49626	Mycobacterium tube
2	1207	97.7	1260	1 V49510	Mycobacterium sp.
3	1207	97.7	1260	1 V49625	Mycobacterium tube
4	455.6	36.9	682	1 V49511	Mycobacterium mari
5	266.8	21.6	1125	1 N91423	Sequence of heat-r
6	202	16.4	28171	1 V52155	Streptococcus pneu
7	168	13.6	9280	1 W74442	Streptococcus aur
8	87.6	7.1	1074	1 T67971	H. pylori membrane
9	69.2	5.6	544	1 V43039	Streptococcus pneu
10	56.8	4.6	31096	1 V74370	Staphylococcus aur
11	50.8	4.1	390	1 Q21833	Randomising oligon
12	50.8	4.1	390	1 Q36859	PCR primer for 5'
13	50.4	4.1	15872	1 T68715	Streptomyces venez
14	46.4	3.8	985	1 V44439	Mycobacterium tube
15	46.4	3.8	985	1 V44548	M. tuberculosis im
16	46	3.7	2582	1 T73117	Actinoplanes sp. a
17	46	3.7	17955	1 V56642	Actinoplanes sp. a
18	45.4	3.7	833	1 Q64203	snab gene encoding
19	45.4	3.7	5392	1 Q64201	Sequence compriain
20	45	3.6	3946	1 T93610	Mycobacterium tube
21	44.8	3.6	535	1 V44428	Mycobacterium tube
22	44.8	3.6	535	1 V44537	M. tuberculosis im
23	44.4	3.6	11495	1 X53481	Human adenosine Al
24	43.6	3.5	28958	1 T66740	Actinoplanes sp. a
25	43.6	3.5	28958	1 T66740	Actinoplanes sp. a
26	43.6	3.5	49377	1 T89562	Scorangium cellulos
27	42.4	3.4	1833	1 V05287	The soraphen biosy
28	42.4	3.4	2185	1 V84066	sbR gene encoding
29	42.4	3.4	2185	1 V09365	Clone p5-4 encodin
30	42.4	3.4	24379	1 T93095	S. lividans protea
31	42.4	3.4	24379	1 T93095	Streptomyces feno
32	42.2	3.4	329	1 V25925	Streptomyces roseo
33	42.2	3.4	329	1 V44425	Mycobacterium tube
34	42	3.4	3946	1 V64534	M. tuberculosis im
				1 T93610	Mycobacterium tube

C	35	42	3.4	4018	1	Q63879	Polyhydroxyalkanoa
	36	41.8	3.4	882	1	V44403	Mycobacterium tube
	37	41.8	3.4	882	1	V64512	M. tuberculosis im
	38	41.6	3.4	1413	1	X34205	Mycobacterium spec
	39	41.6	3.4	1722	1	X34206	Mycobacterium spec
	40	41.6	3.4	12036	1	Q04668	FHA structural gen
	41	41.2	3.3	2414	1	Q05926	Sequence encoding
	42	41	3.3	29879	1	Q46806	eryA region of S.
C	43	40.8	3.3	2151	1	Q76252	HSV-2 protease, IC
C	44	40.8	3.3	2151	1	Q75251	HSV-2 protease/ICP
C	45	40.8	3.3	2472	1	Q84671	HSV-2 UL26 gene. N

ALIGNMENTS

RESULT	1
1	V49626
ID	V49626 standard; DNA; 1245 BP.
AC	V49626;
DT	20-NOV-1998 (first entry)
DE	Mycobacterium tuberculosis L-Alanine dehydrogenase 40 KD antigen.
KM	ss: Alanine dehydrogenase: tuberculosis; TB; mycobacteria; L-alanine.
OS	Mycobacterium tuberculosis.
PN	W09836089-A2.
PD	20-AUG-1998.
PR	29-JAN-1998; E00483.
PR	29-JAN-1997; EP-101338.
PA	(FLOH/) FLOHE L.
PI	Flohe L, Hutter B, Kolk A, Singh M;
DR	WPI: 98-457123/39.
PT	Diagnosis of, e.g. tuberculosis from alanine dehydrogenase activity
PT	- useful for, e.g. for diagnosis, differentiation of strains,
PT	monitoring vaccination and identification of mycobacterial
PT	inhibitors
PS	Disclosure; Fig 3.19; 5'p; German.
CC	The Mycobacterium tuberculosis L-alanine dehydrogenase (AlaDH) is used in
CC	the production of kits for diagnosing tuberculosis (TB) and other
CC	mycobacterial infections in humans or animals. Kits are used for direct
CC	diagnosis of TB on clinical samples (e.g. body fluids) and can
CC	differentiate between pathogenic and non-virulent strains, e.g. for
CC	identifying pathogens of the M. tuberculosis (M. t.) complex. The kit may
CC	also be used to identify substances that inhibit mycobacteria, for
CC	combating epidemics and for vaccination follow-up. Oligonucleotides
CC	derived from AlaDH are used for diagnosis in diagnostic hybridization tests,
CC	also for culture confirmation of isolated strains and for chromosome
CC	fingerprinting to detect/differentiate between mycobacteria, and for
CC	L-alanine-specific biotransformation reactions. AlaDH is an early
CC	antigen, present extracellularly after only a few days of growth, making
CC	it an ideal drug target.
SQ	Sequence 1245 BP; 238 A; 398 C; 387 G; 222 T;

Query Match 97.7%; Score 1207; DB 1; Length 1245;
Best Local Similarity 99.4%; Pred. No. 3.7e-242;
Matches 1234; Conservative 0; Mismatches 1; Indels 7; Gaps 2;

QY	1	ATCTTGCGATTAATCGAACTTTCTTCACACTGAAGGCTACAGATGAGAGGGGTATC	60
DB	1	ATCTTGCGATTAATCGAACTTTCTTCACACTGAAGGCTACAGATGAGAGGGGTATC	60
QY	61	ATGGCGGTGGTATTCGACCGAGCAACCAAAACACG-----AATTCGGGTGGCATC	114
DB	61	ATGGCGGTGGTATTCGACCGAGCAACCAAAACCAAAATTCGAAATTCGGGTGGCATC	120
QY	115	ACCCGCGCGCGGTGGCGGAACTAACCGGTGGTGGCATGAGTGCATCAGCAGGT	174
DB	121	ACCCGCGCGCGGTGGCGGAACTAACCGGTGGTGGCATGAGTGCATCAGCAGGT	180
QY	175	GCCGAGAGAGGGGTGGCGGTATCAGCGAGCGGATTTCAAGCGGCGGCGCAATGTC	234
DB	181	GCCGAGAGAGGGGTGGCGGTATCAGCGAGCGGATTTCAAGCGGCGGCGCAATGTC	240

```

QY 235 GGCACCCGCGACAGGTGTGGGCGGACGCTGATTTATTTGCTCAAGAGTCAAGAACCGATA 294
    |||||
DB 241 GGCACCGCGCGACAGGTGTGGGCGGACGCTGATTTATTTGCTCAAGAGTCAAGAACCGATA 300
QY 295 GCGGCGGAATAGCGGCGGCTCGGACACGGGC-GATCTTGTCTACAGTTCTTGCATTTGGCC 353
    |||||
DB 301 GCGGCGGAATAGCGGCGGCTCGGACACGGGC-GATCTTGTCTACAGTTCTTGCATTTGGCC 360
QY 354 GCGTCACGTGCTTCACCGATGCTGTTGGATTCGCGCACACGCTCAATTGCTTACGAG 413
    |||||
DB 361 GCGTCACGTGCTTCACCGATGCTGTTGGATTCGCGCACACGCTCAATTGCTTACGAG 420
QY 414 ACCGTCCAGACCGCGGACGCGGCTACCTCTGCTTGGCCCGATGAGCGAAATGCGCGGT 473
    |||||
DB 421 ACCGTCCAGACCGCGGACGCGGCTACCTCTGCTTGGCCCGATGAGCGAAATGCGCGGT 480
QY 474 CGACTCGCGCGCGAGGTTGGCGCTTACCACTGATGCGAACCCAAAGGGGGCGGCGGTG 533
    |||||
DB 481 CGACTCGCGCGCGAGGTTGGCGCTTACCACTGATGCGAACCCAAAGGGGGCGGCGGTG 540
QY 534 CTGATGGGCGGGGTGCGCGGCGCTGAAACCGGCGGACGTCG-GGTGATCGGCGCGGACCC 593
    |||||
DB 541 CTGATGGGCGGGGTGCGCGGCGCTGAAACCGGCGGACGTCG-GGTGATCGGCGCGGACCC 600
QY 594 GCGGCGTACAAAGGACGCGGCGATGCGCAAGCGCATGGGCGGACGCTTACGCTTACGAC 653
    |||||
DB 601 GCGGCGTACAAAGGACGCGGCGATGCGCAAGCGCATGGGCGGACGCTTACGCTTACGAC 660
QY 654 ATCAACATCGACAAATTTGCGCAACTGACGCGGAGTTCTGCGCGCGGATCCACACTCGC 713
    |||||
DB 661 ATCAACATCGACAAATTTGCGCAACTGACGCGGAGTTCTGCGCGCGGATCCACACTCGC 720
QY 714 TACATCAGGCTGACGAGCTGAGAGGTGCGGTCAAAAGCGCGGACGCTGATGAGGCGCC 773
    |||||
DB 721 TACATCAGGCTGACGAGCTGAGAGGTGCGGTCAAAAGCGCGGACGCTGATGAGGCGCC 780
QY 774 GTCTGTGTCGAGGCGGCGCAAGGACCCAAATTAATCTGCAATTCATCTGTCGCGCATATG 833
    |||||
DB 781 GTCTGTGTCGAGGCGGCGCAAGGACCCAAATTAATCTGCAATTCATCTGTCGCGCATATG 840
QY 834 AAACAGAGTGGGTGATCTGTGTGATATAGCCATGACACGAGGCGGCTGTTTCGAAGGCTCA 893
    |||||
DB 841 AAACAGAGTGGGTGATCTGTGTGATATAGCCATGACACGAGGCGGCTGTTTCGAAGGCTCA 900
QY 894 CGACGACGACCTACGACCAACCCGACGTTGCGCGTGCACGCAAGCTGTTTACTGCGTG 953
    |||||
DB 901 CGACGACGACCTACGACCAACCCGACGTTGCGCGTGCACGCAAGCTGTTTACTGCGTG 960
QY 954 GCGAATATGCGCGCTGCGGTGCGGAGACGTCGACTACGCGCTGACCAACGCGGAGATG 1013
    |||||
DB 961 GCGAATATGCGCGCTGCGGTGCGGAGACGTCGACTACGCGCTGACCAACGCGGAGATG 1020
QY 1014 CCGATATGCTGAGGCTTGGCGACCATGGCTGGGCGGCGGCGGTGCGGTGATCCGCGCA 1073
    |||||
DB 1021 CCGATATGCTGAGGCTTGGCGACCATGGCTGGGCGGCGGCGGTGCGGTGATCCGCGCA 1080
QY 1074 CTAGCCAAAGGTCTTTTGCAGCGACGAGAGGGGCGTTACTGTCTCCGAAACGGGTGCGCACCGAC 1133
    |||||
DB 1081 CTAGCCAAAGGTCTTTTGCAGCGACGAGAGGGGCGTTACTGTCTCCGAAACGGGTGCGCACCGAC 1140
QY 1134 CTGGGGGTGCGGCTTACCGAGGCGCGGCGGAGGCTGTGCTGCTGAGCTCTTGGCGGCTGTTAC 1193
    |||||
DB 1141 CTGGGGGTGCGGCTTACCGAGGCGCGGCGGAGGCTGTGCTGCTGAGCTCTTGGCGGCTGTTAC 1200
QY 1194 GCGGAGACACGTCGCGGAGTAAAGGAGAGCATGATGTCGCGC 1235
    |||||
DB 1201 GCGGAGACACGTCGCGGAGTAAAGGAGAGCATGATGTCGCGC 1242

```

RESULT 2
V49510
ID V49510 standard; DNA; 1260 BP.
AC V49510;

```

DT 20-OCT-1998 (first entry)
DE Mycobacterium sp. Aladh DNA.
KW Alanine dehydrogenase; Aladh; ADH; diagnosis; tuberculosis; pathogen;
    swimmers disease; vaccine; epidemic; infection; identification; ss.
OS Mycobacterium sp.
PN M09832862-A2.
PF 30-JUL-1998.
PR 29-JAN-1998; E00484.
PS (FLOH/) FLOHE L.
PI FLOHE L. Hutter B. Kolk A. Singh M.
DR MPI: 98-427958/36.
PT Nucleic acid encoding alanine dehydrogenase of Mycobacterium marinum
    - used for diagnosis of tuberculosis and other mycobacterial
    diseases, also for treatment and prevention, for drug screening and
    PT for bio-transformation
PS disclosure, page 11; 57pp; German.
CC This sequence encodes an alanine dehydrogenase (ADH) protein isolated
    from a Mycobacterium sp. This protein is used to diagnose tuberculosis
    and other mycobacterial infections (including 'swimmers disease', caused
    by M. marinum, a fish pathogen) in humans or animals. The protein can
    also be used for control of epidemics and for vaccination, to screen for
    CC agents with anti-mycobacterial activity, and in bio-transformations that
    are specific for L-alanine. Also mycobacteria can be identified by
    CC analysis of genomic ADH sequences. ADH is an antigen that is secreted
    early during infection.
SQ Sequence 1260 BP; 243 A; 403 C; 389 G; 225 T;

Query Match 97.7%; Score 1207; DB 1; Length 1260;
Best local similarity 99.4%; Pred. No. 3.7e-242;
Matches 1234; Conservative 0; Mismatches 1; Indels 7; Gaps 2;

QY 1 ATCTTCAGATTAATGAACTTTCTTACACCTGAAAGCGTACAGTATCGAGAGGGTAAATC 60
    |||||
DB 16 ATCTTCAGATTAATGAACTTTCTTACACCTGAAAGCGTACAGTATCGAGAGGGTAAATC 75
QY 61 ATCCGCGTGGATATTCGCA :GAGACAAAACAAACG-----AATCCGGGTGGCATC 114
    |||||
DB 76 ATCCGCGTGGATATTCGCA :GAGACAAAACAAACG-----AATCCGGGTGGCATC 135
QY 115 ACCCGCGCGGCGTGGCGGACCACTACCGGCTGAGGCGATGAGTGCATCCAGGCGAGT 174
    |||||
DB 136 ACCCGCGCGGCGTGGCGGACCACTACCGGCTGAGGCGATGAGTGCATCCAGGCGAGT 195
QY 175 GCGGAGAGGGCTCGGCTATTCACACGCGGATTTTCAGGCGGCGGCGGCACTGCTC 234
    |||||
DB 196 GCGGAGAGGGCTCGGCTATTCACACGCGGATTTTCAGGCGGCGGCGGCACTGCTC 255
QY 235 GCGACGCGGACGAGTGTGGGCGGCGGACGCTGATTTATTTGCTCAAGAGTCAAGAACCGATA 294
    |||||
DB 256 GCGACGCGGACGAGTGTGGGCGGCGGACGCTGATTTATTTGCTCAAGAGTCAAGAACCGATA 315
QY 295 GCGGCGGAATAGCGGCGGCTCGGACACGGGC-GATCTTGTCTACAGTTCTTGCATTTGGCC 353
    |||||
DB 316 GCGGCGGAATAGCGGCGGCTCGGACACGGGCAGATCTTGTACGCTTCTTGATTTGGCC 375
QY 354 GCGTACGTGCTTGCACGATGCTGTTGGATTTCCGCGACCAAGTCAATTGCTTACGAG 413
    |||||
DB 376 GCGTACGTGCTTGCACGATGCTGTTGGATTTCCGCGACCAAGTCAATTGCTTACGAG 435
QY 414 ACCGTCCAGACCGCGGACGCGGCTACCTCTGCTTGGCCCGATGAGCGAAATGCGCGGT 473
    |||||
DB 436 ACCGTCCAGACCGCGGACGCGGCTACCTCTGCTTGGCCCGATGAGCGAAATGCGCGGT 495
QY 474 CGACTGCGCGCGGAGTGGCGCTTACCACTGATGCGAAACCAAGGGGGCGGCGGTG 533
    |||||
DB 496 CGACTGCGCGCGGAGTGGCGCTTACCACTGATGCGAAACCAAGGGGGCGGCGGTG 555
QY 534 CTGATGGGCGGGGTGCGGCGGCTGCAACCGGCGGACGCTGATGATGCGGCGGCGGAC 593
    |||||
DB 556 CTGATGGGCGGGGTGCGGCGGCTGCAACCGGCGGACGCTGATGATGCGGCGGCGGAC 615

```

OY		594	GCGGGCTACAAAGGAGGCCGCATTCGGCAAGGCGATGGGGCGAACCGTTTACAG	653
Db		616	GCGGGCTACAACGACGAGCCGCATTCGCCMAAGCGCATGGGGCGAACCGTTTACAG	675
OY		654	ATCAACATGACAAAATTCTGGCACTCGAAGCGCGAATTCTGGGGCGAATCCACTGCG	713
Db		676	ATCAACATGACAAAATTCTGGCAACTCGAAGCGCGAATTCTGGGGCGAATCCACTGCG	735
OY		714	TATCATTACGGGCTACAGGCTCGAGGGTGCCGCTTAACCTGCCGACTGTGATTGGGGCC	773
Db		736	TATCATTACGGGCTACAGGCTCGAGGGTGCCGCTTAACCTGCCGACTGTGATTGGGGCC	795
OY		774	GTCCTGGTGGCAGGCGCGACGAAGCACCAATAATGATCTCAATTCACCTGTCCGCATATG	833
Db		796	GTCCTGGTGGCAGGCGCGACGAAGCACCAATAATGATCTCAATTCACCTGTCCGCATATG	855
OY		834	AACCCAGGTGGCGGTACTGTGTGATATAGCCATCGACCAGGGCGGCTGTTTGAAAGCTCA	893
Db		856	AACCCAGGTGGCGGTACTGTGTGATATAGCCATCGACCAGGGCGGCTGTTTGAAAGCTCA	915
OY		894	CGACCGACCAACCTACGACCAACCGGAGTTGGCCGTGACAGACGCTGTTTACATGCGTG	953
Db		916	CGACCGACCAACCTACGACCAACCGGAGTTGGCCGTGACAGACGCTGTTTACATGCGTG	975
OY		954	GCGAATATGCCCCGCTCGGTGTCGCCGAAGCGTCGACCTACGCGCTGACCAACCGCAGATG	1013
Db		976	GCGAATATGCCCCGCTCGGTGTCGCCGAAGCGTCGACCTACGCGCTGACCAACCGCAGATG	1035
OY		1014	CCGATATGCTCGAGGCTTGGCGAACATAGCTGGCGGGCGGCTGGCGGTGCAATCCGGCA	1073
Db		1036	CCGATATGCTCGAGGCTTGGCGAACATAGCTGGCGGGCGGCTGGCGGTGCAATCCGGCA	1095
OY		1074	CTACCCCAAAGGTCTTTTGCAGCGCACAGAAGGGGCTTACTCTCCGAACGGGTGGCCACAC	1133
Db		1096	CTACCCCAAAGGTCTTTTGCAGCGCACAGAAGGGGCTTACTCTCCGAACGGGTGGCCACAC	1155
OY		1134	CTGGGGGTGCCGTTACCCGAGCCCAGCGTAGCTGTGGCTGACTCTGGCGCGCTGTTAC	1193
Db		1156	CTGGGGGTGCCGTTACCCGAGCCCAGCGTAGCTGTGGCTGACTCTGGCGCGCTGTTAC	1215
OY		1194	GCGANCAACAGTGGGAGTAAAGGAAGCGATGATCGGGC	1235
Db		1216	GCGANCAACAGTGGGAGTAAAGGAAGCGATGATCGGGC	1257
RESULT		3		
V49625		V49625	standard; DNA; 1260 BP.	
AC		V49625:		
DT		20-NOV-1998	(first entry)	
DE		Mycobacterium tuberculosis	Alanine dehydrogenase.	
KW		ss; Alanine dehydrogenase;	tuberculosis; TB; mycobacteria; L-alanine.	
OS		Mycobacterium tuberculosis.		
PN		MO9836089-AZ.		
PD		20-AUG-1998		
PF		29-JAN-1998:	E00483	
PR		29-JAN-1997:	EP-101338.	
PA		(FLOH/) FLOHE L.		
PI		Flohe L., Hutter B., Kolik A., Singh M,		
DR		WPI: 98-457123/39.		
PT		Diagnosis of, e.g. tuberculosis from alanine dehydrogenase activity		
PT		- useful for, e.g. for diagnosis, differentiation of strains,		
PT		monitoring vaccination and identification of mycobacterial		
PT		inhibitors		
PS		Claim 13: Fig 2.3; 55pp; German.		
CC		The Mycobacterium tuberculosis alanine dehydrogenase (AlaDH) is in		
CC		the production of kits for diagnosing tuberculosis (TB) and other		
CC		mycobacterial infections in humans or animals. Kits are used for direct		
CC		diagnosis of TB on clinical samples (e.g. body fluids) and can		
CC		differentiate between pathogenic and non-virulent strains, e.g. for		
CC		identifying pathogens of the M.tuberculosis (M. t.) complex. The kit may		
CC		also be used to identify substances that inhibit mycobacteria, for		
CC		combating epidemics and for vaccination follw-up. Oligonucleotides		

CC derived from AlAdh are used similarly in diagnostic hybridisation tests, CC also for culture confirmation of isolated strains and for chromosome CC fingerprinting to detect/differentiate between mycobacteria, and for CC L-alanine-specific biotransformation reactions. AlAdh is an early CC antigen, present extracellularly after only a few days of growth, making CC it an ideal drug target.

Sequence 1260 BP: 243 A; 403 C; 389 G; 225 T;

Query Match	97.7%	Score 1207;	DB 1;	Length 1260;
Best Local Similarity	99.4%	Pred. No. 3.7e-242;		
Matches 1234;	Conservative	0;	Mismatches 1;	Indels 7; Gaps 2

QY	1	ATCTCGAGATTATTCGAACCTTCTTCAOACTGAAGGTACAGTATCGAAGGGGTATC	60
Db	16	ATCTTCGAGATTATTCCTAATCTTCTTCACTGAAGGTACAGTATCGAAGGGGTATC	75
QY	61	ATGCGCGTGGTATTTCGACCGAGACCMAAACMACG-----AATTCGGGTGGCCATC	114
Db	76	ATGCGCGTGGTATTTCGACCGAGACCMAAACMAAGAAATTCGAATTTCGGGGTGGCCATC	135
QY	115	ACCCCGCGCCGGCTTCGGGAACTAACCCGTCGTGGCCATGAGTGGCTCATTCACAGCAGGT	174
Db	136	ACCCCGCGCCGGCTTCGGGAACTAACCCGTCGTGGCCATGAGTGGCTCATTCACAGCAGGT	195
QY	175	GCCGGAGAGGGCTGGGCTATACCGACGCGGATTCCTAAGCGGGGACGGCAGCAACTGGTC	234
Db	196	GCCGGAGAGGGCTGGGCTATTCACCGACGCGGATTCCTAAGCGGGGACGGCAGCAACTGGTC	255
QY	235	GGCACCCCGCAGCAGGTGTGGCGCGAGCTGATTTATTGCTCAAGTCAAAAGAACCGATA	294
Db	256	GGCACCCCGCAGCAGGTGTGGCGCGAGCTGATTTATTGCTCAAGTCAAAAGAACCGATA	315
QY	295	GCGGGGGAATGAGGGCCGGCTTCGGACAGCGGC-GATCTTTGTTACAGTCTTGCAATTGGCC	353
Db	316	GCGGGGGAATGAGGGCCGGCTTCGGACAGCGGC-GATCTTTGTTACAGTCTTGCAATTGGCC	375
QY	354	GCGTCACGTGCTTCGACCGATGACCTTGTGGATTCCGGGACACACGTCAATTGGCTACGAG	413
Db	376	GCGTCACGTGCTTCGACCGATGACCGATGACCTTGTGGATTCCGGGACACACGTCAATTGGCTACGAG	435
QY	414	ACCGTCCAGACCCCGCCGACGGCGCACTACCCCTGCTCCCGCGATGAGCAGAAATGCCGGT	473
Db	436	ACCGTCCAGACCCCGCCGACGGCGCACTACCCCTGCTCCCGCGATGAGCAGAAATGCCGGT	495
QY	474	CGACTCGCGCGCCGAGGTGCGCTCTTACCACTGATGGAAACCCAAAGGGGGCCGGGTGTG	533
Db	496	CGACTCGCGCGCCGAGGTGCGCTCTTACCACTGATGGAAACCCAAAGGGGGCCGGGTGTG	555
QY	534	CTGATGGCGGGGTGCCCGCGCGGTGGAACCGGCCACAGCTCGTGTGTATCGCGCGCGCAC	593
Db	556	CTGATGGCGGGGTGCCCGCGCGGTGGAACCGGCCACAGCTCGTGTGTATCGCGCGCGCAC	615
QY	594	GCGCGCTACAAACGACGCCCGCATGCGCAAGCGCATGGGCGCGACCGTTACGGTTCTAGAC	653
Db	616	GCGCGCTACAAACGACGCCCGCATGCGCAAGCGCATGGGCGCGACCGTTACGGTTCTAGAC	675
QY	654	ATCAACATCGACAACCTTCCGCACTCGACGCCGAGTTTGGCGGCGGGATCGACACTGGC	713
Db	676	ATCAACATCGACAACCTTCCGCACTCGACGCCGAGTTTGGCGGCGGGATCGACACTGGC	735
QY	714	TACTCATGGGCTCTAGAGCTGGAAGGTGGCGCTCAAAAGTGGCGGACCTGGGTGATTGGGGCC	773
Db	736	TACTCATGGGCTCTAGAGCTGGAAGGTGGCGCTCAAAAGTGGCGGACCTGGGTGATTGGGGCC	795
QY	774	GTCCTGGTGCAGGCGCCCAAGGACCCAAATTAAGTCTCGAATTACTGTGACGGCATATG	833
Db	796	GTCCTGGTGCAGGCGCCCAAGGACCCAAATTAAGTCTCGAATTACTGTGACGGCATATG	855
QY	834	AAACCAAGGTGGGTAACGTGGTGAATATAGCATGCAACAGGCGCGCTTTTCGAAGGCTCA	893
Db	856	AAACCAAGGTGGGTAACGTGGTGAATATAGCATGCAACAGGCGCGCTTTTCGAAGGCTCA	915

QY 894 CGACCGACACCTACGACACCGACGCTGCGGTCAGACACAGCTTTACTGCGT 953
 DB 916 CGACCGACACCTACGACACCGACGCTGCGGTCAGACACAGCTTTACTGCGT 975
 QY 954 GCGAACATGCGCGCGCGTGGTCCGAGACGTCGACCTACGCGCTGACCAACGCGAGATG 1013
 DB 976 GCGAACATGCGCGCGCGTGGTCCGAGACGTCGACCTACGCGCTGACCAACGCGAGATG 1035
 QY 1014 CCGATATGCTCGACGCTTGGCGACGACGATGGTGGGCGCGCGTGGTGGATCCGCGCA 1073
 DB 1036 CCGATATGCTCGACGCTTGGCGACGACGATGGTGGGCGCGCGTGGTGGATCCGCGCA 1095
 QY 1074 CTACCCAAAGGTCTTCTTTCGACGACGAGAGGCGCTTACTGTCCGAAAGGATGGCCACCGAC 1133
 DB 1096 CTACCCAAAGGTCTTCTTTCGACGACGAGAGGCGCTTACTGTCCGAAAGGATGGCCACCGAC 1155
 QY 1134 CTGGGGGTGCGCTTACCGACGACCGCGCGCGCGCGCGCGCTGACCTGCGCGCGCTTAC 1193
 DB 1156 CTGGGGGTGCGCTTACCGACGACCGCGCGCGCGCGCGCGCTGACCTGCGCGCGCTTAC 1215
 QY 1194 GCGGACACAGCTGCGGAGTAAGGAGGAGCGATGATGTCGCGC 1235
 DB 1216 GCGGACACAGCTGCGGAGTAAGGAGGAGCGATGATGTCGCGC 1257

RESULT 4

ID V49511 standard; DNA; 682 BP.
 AC V49511;
 DT 20-OCT-1998 (first entry)
 DE Mycobacterium marinum Mar3 DNA.
 KW Alanine dehydrogenase; AlADH; ADH; diagnosis; tuberculosis; pathogen;
 SW swimmers disease; vaccine; epidemic; infection; identification; ss.
 OS Mycobacterium marinum.
 FH Key Location/Qualifiers
 FT CDS 1..682
 FT /*tag= a
 FT /codon_start= 3
 FT /product= "Mar3"
 FT /note= "Alanine dehydrogenase"
 PN W09832862-A2.
 PD 30-JUL-1998.
 PE 29-JAN-1998; E00484.
 PR 29-JAN-1997; EP-101339.
 PA (FIORE) FIORE L, Kolk A, Singh M;
 PI FIORE L, Butler B, Kolk A, Singh M;
 DR WPI: 98-427958/36.
 DR P-PSDB: W64481.
 PT Nucleic acid encoding alanine dehydrogenase of Mycobacterium marinum
 PT - used for diagnosis of tuberculosis and other mycobacterial
 PT diseases, also for treatment and prevention, for drug screening and
 PT for bio-transformation
 PS Claim 1; Page 34-35; 57pp; German.
 CC This sequence encodes an alanine dehydrogenase (ADH) protein, Mar3
 CC isolated from Mycobacterium marinum. This protein is used to diagnose
 CC tuberculosis and other mycobacterial infections (including 'swimmers'
 CC disease', caused by M. marinum, a fish pathogen) in humans or animals.
 CC The protein can also be used for control of epidemics and for
 CC vaccination, to screen for agents with anti-mycobacterial activity, and
 CC in bio-transformations that are specific for L-alanine. Also mycobacteria
 CC can be identified by analysis of genomic ADH sequences. ADH is an antigen
 CC that is secreted early during infection.
 SQ Sequence 682 BP; 105 A; 254 C; 225 G; 98 T;

Query Match 36.9%; Score 455.6; DB 1; Length 682;
 Best Local Similarity 80.2%; Pred. No. 2.1e-86;
 Matches 547; Conservative 0; Mismatches 134; Indels 1; Gaps 1;

QY 98 AATTCGGGTGCGCATCCCGCGCGTGCAGACTAACCGCTGCGCATGAGG 157
 DB 1 AATTCGGGTGCGCATCCCGCGCGTGCAGACTAACCGCTGCGCATGAGG 60

QY 158 TCGTCATCCAGGACAGTCCGCGAGAGGCTCGCTATCACCGACGCGATTTCAGGCGG 217
 DB 61 TCGTCATCCAGGACAGTCCGCGAGAGGCTCGCTATCCAGGACGCGATTTCAGGCGG 120
 QY 218 CAGGCGCGCACTGCTGGCACCCCGACAGGTGTGGGCGGACCTGATTATGCTCA 277
 DB 121 CCGGTGCGCGGCTGATACACCGCGGACAGGTGTGGGCGGACCTGATTATGCTCA 180
 QY 278 AGGTCAAGAACCGATAGAGGCGGGAATACAGGCGGCTGGACAGCGAGG- GATCTGTCTCA 336
 DB 181 AGGTCAAGAACCGATAGAGTCCGAGTACGCGGCTGGCGGCGGACAGCTGCTTCA 240
 QY 337 CGTTCTTGGATTGCGCGCGGTGACGCTGCTTGCACCGATGCGTTGGATTCCGCGACCA 396
 DB 241 CTTACCTGCACTGCGCGGCTGCGCGGCTGCGACCGATGCGTTGGATTCCGCGACCA 300
 QY 397 CGTCAATTCGCTACGACCGCTGACAGCGCGGCGGCGGACGATGCGTTGGATTCCGCGACCA 456
 DB 301 CGTCAATTCGCTACGAGCGGTGACAGCGCGGCGGCGGACGATGCGTTGGATTCCGCGACCA 360
 QY 457 TGAGCGAAGTCCGCGGCTGACCTGCGCGCGGCGGCGGCTTACGACCTGATGCGACCC 516
 DB 361 TGAGCGAAGTCCGCGGCGGCTGCGCGCGGCGGCGGCTTACGACCTGATGCGACCC 420
 QY 517 AAGGCGCGCGGTGCTGATGCGCGGCTGCGCGGCTGCGACCGCGGCGGCGGCGG 576
 DB 421 ACGGCGCGCGGTGCTGATGCGCGGCTGCGCGGCTGCGACCGCTGCGGCGGCGG 480
 QY 577 TGATCGGCGCGGCGGCGGCGGCTGACGACGAGCGGCGGCGGCGGCGGCGGCGG 636
 DB 481 TGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
 QY 637 CCGTTACGCTTACGATACATGACATGACAACTTGGCACTGACCGCGGAGTTGCG 696
 DB 541 TGTCTACCGGCTGATGATGATCAATCAATCAATGCGCGGCGGAGTTGCGG 600
 QY 697 GCGGATCCACACTCGCTACTGCTGCGGCTGACGAGCTGCGGAGGTCGCTCAACAGTGGCG 756
 DB 601 GTGCGCTGCGGACCGGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 QY 757 ACCTGCTGATGGGCGGCTCT 778
 DB 661 ACATGCTGATGGGCGGCTCT 682

RESULT 5

ID N91423 standard; DNA; 1125 BP.
 AC N91423;
 DT 01-FEB-1991 (first entry)
 DE Sequence of heat-resistant alanine dehydrogenase (AH) gene with mol. wt.
 DE 2 MD or less
 KW Enzyme; ds.
 OS Bacillus stearothermophilus IFO 12550.
 PN J01043194-A.
 PE 15-FEB-1989.
 PE 10-AUG-1987; 200524.
 PR 10-AUG-1987; JP-200524.
 PA (NIRA) Unilika KK.
 DR WPI: 89-096096/13.
 PT Recombinant plasmid for transforming Escherichia coli -
 PT obt. by connecting heat resistant alanine dehydrogenase gene to
 PT vector plasmid
 PS Disclosure: Fig 2, p63; 9pp; Japanese.
 CC A recombinant plasmid contg. heat-resistant alanine dehydrogenase (AH)
 CC gene with mol. wt. 2 or less MD is claimed. Cells transformed with the
 CC vector produce high levels of heat-resistant AH. Transformed E. coli
 CC is useful as a clinical inspection reagent.
 SQ Sequence 1125 BP; 246 A; 295 C; 367 G; 217 T;

Query Match 21.6%; Score 266.8; DB 1; Length 1125;
 Best Local Sim. larity 56.2%; Pred. No. 3e-47;

Matches 604; Conservative 0; Mismatches 457; Indels 13; Gaps 5;

```

QY 61 ATGCGCGTGGTATTCGACCGAGACCAAAACAAATTCGGGGTGGCCATCACC 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 ATGAGATCGGATTCGACCAAAATCAAAATCAAAATCAAAATCAAAATCA 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 GCGGCGTGGGAAATACCCGTCGTGGCCATGATGCTGCT---CATCGAGCGAGTGC 177
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 GCGAGCGCTGATGAGCGCTCGTCAAAAGCGGGCATGAGTGTGATGAGAGCGAAGGC 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 178 GGAAGAGCGCTGCTATCACCAGCGGATTTCAAGCGGCGGCGGCACTGCTGCG 237
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 GCGGCTGGTGGGGTTCCTCGATTCGATATGAAAGAACCGGCGAGCTGATGCTTGC 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 238 ACCGCGCAGCAGTGTGGGCGGAGCGTATGCTCAAGGTCGAAGAACGATGCG 247
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 CGAAGCTGAGAGATGCTTGGAGCGGAGATGCTGTTGAAGTGAAGAGCGCGTGC 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 298 GCGGAATACGCGCGCTGCGACACAGG-GCGATCTTGTTCAGTCTCTCATTTGGCGCG 356
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 CGAGAGTCCGCTATTTCCCGCGATGATTTGTTGATGCTATTTGCTATTAAGCGCG 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 357 TCAGTGTCTGACCGATGCGTGTGATTCGCGACACGATGCTATGCTACGAGACC 416
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 GCGGAACGCTGACGAAAGCGCTGCTGAGCAAAAGTGGTGGCATGCTTACGAGACG 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 417 GTCCAGACCCCGACGGCGCATACCTGCTTCCCGCATGAGCGAAGTCGCG 476
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 GTGAGTGTGCGAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 477 CTGCGCCCGCAGGTTGGGCTTACCATGATGAGCAACCAAGGGGCGCGGTGCTG 536
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 421 ATGTGCTGCAATGCTGCGCGCGCATTTCTGAGAAAGCGGAGGAGGAGGATTTTG 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 537 ATGAGCGGGTCCCGCGCTGCAACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 596
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 481 CTGAGCGGCTGCGCGGAGTGGCGCGCGGCAAAATGACATGATGCGCGGCGGAGCGG 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 597 GCGTACAGCCAGCCCGCATGCGCAAGGATGGGCGCGCATGCTGATGCTTACGATC 656
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 541 GGGAGCAGACCGGCGAAATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 657 AACATGACAACTGCGCAACTGCGCACTGCGCACTGCGCACTGCGCACTGCGTAC 716
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 601 AAGCGCGAGCGGCTGCGCGCGCATGATGATTTGTCGCGCACCGCTGACGAGCGCTATG 660
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 717 TCATCGGCTGACGCTGAGAGGTGCGCTCAAAAGCTGCTGATGCTGAGTGGCGCTG 776
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 661 TCCAACTGATACATATGCGCGAGTGGCTGCGCAATCGGATTTGCTGCTGCTGCTG 720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 777 CTGCTGCGACGGCGCAAGGACCAATAGTCTGCAATTCATCTGCTGCGCATATGAAA 836
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 721 TTGATCCCGGGGCGGAA---AGCGAAGCTGCTGAGCGAGAGAGATGCTGATACG 777
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 837 CCAAGTGGTACGCTGAGATATAGCATGAGACGAGGCGCTGCTGCAAA---GGCTCA 893
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 778 CCGGAGATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 837
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 894 GCAACGACCACTAGACCAACCGGAGCTGCTGCAAGACAGCTGCTGCTGCTGCTG 953
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 838 CGACCGTACGACGAGTCCGACATAGCTCAAGCAGCGCGCTGCTGCTGCTGCTGCTG 897
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 954 GCGAATGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1013
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 898 GCGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 954
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1014 CCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1073
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 955 CCAATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1014
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1074 CTAGCCAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1127
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1015 CTGTTAAAGGATCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1068
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 6

V52155/C

V52155 standard; DNA; 28171 BP.

ID V52155;

AC 23-OCT-1998 (first entry)

DE Streptococcus pneumoniae genome fragment SEQ ID NO:22.

KW Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;

KM computer readable medium; vaccine; pharmaceutical composition; ds.

PN Streptococcus pneumoniae.

PD MO9818931-A2.

PN 07-MAY-1998.

PF 30-OCT-1997; U19588.

PR 31-OCT-1996; US-029960.

PA (HUMAN) HUMAN GENOME SCI INC.

PI Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M,

PI Kunsch CA, Rosen CA;

DR MPI: 98-27225/24.

PT Computer-readable medium with recorded Streptococcus pneumoniae

PT polynucleotide sequences - useful in diagnostic kits and assays, and

PT pharmaceutical compositions and vaccines for Streptococcus

PT pneumoniae.

PS Claim 1; Page 273-289; 140992; English.

CC The present invention describes a computer readable medium which has

CC the nucleotide sequences SEQ ID NO:1 to 391 (V52134 to V52524) recorded

CC on it, or a representative fragment or a sequence at least 95% identical

CC to SEQ ID NO:1 to 391. The nucleotide sequences depicted in SEQ ID NO:1

CC to 391 (V52134 to V52524) are genomic fragments from Streptococcus

CC pneumoniae. The present invention also describes an isolated nucleic acid

CC molecule encoding a homologue of any of the fragments of the S. pneumoniae

CC genome (SEQ ID NO:1 to 391) where the nucleic acid molecule is produced

CC by a process comprising: (a) screening a genomic DNA library using as a

CC probe a target sequence defined by any of the sequences in SEQ ID NO:1

CC to 391, identifying members of the library which contain sequences

CC that hybridize to the target sequence and isolating the nucleic acid

CC molecules from the members; or (b) isolating mRNA, DNA or cDNA produced

CC from an organism, amplifying nucleic acid molecules whose nucleotide

CC sequence is homologous to amplification primers derived from the

CC fragment of the S. pneumoniae genome to prime the amplification and

CC isolating the amplified sequences. The computer readable medium can be

CC used in a computer-based system for identifying fragments of the

CC S. pneumoniae genome of commercial importance, or expression modulating

CC fragments of the S. pneumoniae genome. Products from the present

CC invention can be used in diagnostic kits and assays, and pharmaceutical

CC compositions and vaccines for S. pneumoniae.

CC Sequence 28171 BP; 9507 A; 5080 C; 6440 G; 8144 T;

Query Match 16.4%; Score 202; DB 1; Length 28171;

Best Local Similarity 53.9%; Pred. No. 16-33;

Matches 533; Conservative 0; Mismatches 435; Indels 20; Gaps 5;

```

QY 61 ATGCGCGTGGTATTCGACCGAGACCAAAACAAAGATTCGGGGTGGCCATCACC 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21711 ATGTTATCGGAATCCCAAAAGAAATTAATAAGAAAACCGTCCCTCAACCT 21652
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 GCGGCGTGGGAAATACCCGTCGTGGCCATGATGCTGCT---CATCGAGCGAGTGC 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21651 GCAAGTGTTCATAGCTAGTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 21592
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 GAGGCGTGGCTGACCAACCGGAGTTTCAGCGCGACCGCGCACTGCTGCTGCTG 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21591 CTGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 21532
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 GCGGACGAGTGTGGCGCGAGCTGATTTATGCTCAAGGTCGAAGAACGATAGCGCG 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21531 GCTGTTAAAGCTTGGG---AGCAGAGTGTGTTGAAGTAAAGATCTTAAGTTCT 21475
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 301 GAATACGCGCGCTGCGACCGGCG---GATCTGTTACAGTCTTTCGATTTGGCGGCTCA 359
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21474 GAATACGCTTACTTGGCGCGAGTCTTCTTCTTACCACTTGCACATGCGCGCTGCT 21415
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```

OY 360 CGTGTCTGACCGATCGTGTGTGATTCGGCAGCCAGCTCAATTGCTACGACCGCTC 419
DB 21414 CCAGATTTAGCAGATCTGTT-----AACAGCAAAAACACTGAACTGTT 21367
OY 420 CAGACCCCGGACGCGCTACTACCCCTGCTTGGCCCGACGAGCAAGTCCGCGTACTC 479
DB 21366 CGTGAACATCAAGAGCAACATCCGCTCTCTGTTCCATGATGAGTGCAGTGTATG 21307
OY 480 GCGCGCCAGTGGCGCTGATCCACCTGATGCCAACCAGGGGGCGCGCTGCTGATG 539
DB 21306 GCGTTCATATGAGCTACTCTCTACTAGCAAGCTGGTGGCTGCTGTTCTACTT 21247
OY 540 GCGCGGCTCGCGCTCGCAACCGCGGAGCTGCTGATGGCGCGCGCCAGCCGCGC 599
DB 21246 GCGTGTGTCAGAGTGTTCCTCAAAAGGAAAGTAACTATCATGCTGCTGCTCGCT 21187
OY 600 TACACGCAAGCCCGATGCGCAACGCGATGGCGGCGACCGTTCAGATCAAC 659
DB 21186 ACACATGCTGCGCGCTCCCTGCTGCTGCTCAAGTACTATTTATGATTTAGT 21127
OY 660 ATCGACAACCTGCGCACTCGACCGCGAGTCTGCGCGGCGGATCCACACTGCTACTCA 719
DB 21126 TCGAACGCTTCTCACTTATAGAGAGTCTTGGAGATTAATTCAACTCTTATGCT 21067
OY 720 TCGGCTACGAGCTGAGGCTGCGCTCAACGCTGCGACCTGCTGATGGCGCGCTCTG 779
DB 21066 AATTCATTAACATTAAGCAAGTGTGAGAGATGCTGATGTGATGGACCATTC 21007
OY 780 GTGCCAGGCGCCAGGACCAATTAATGCTGATTCCTGTGCGCATATGAACCA 839
DB 21006 ATCCCTGGGCAAAACACCGGATTTGGAGAGATGAGATGCAAAACAAATGGCTCA 20947
OY 840 GGTGGGCTGCTGCTGATTAAGCATCGACAGGCGGCGTGTTCGA---AGGCTACGA 396
DB 20946 GCGCTCTGTA-TGTTTACCTTGTCTGTGACCAAGTGGCGCTTATGAAACACTGACCT 20888
OY 897 CGGACCACTACGACACCGGAGCTTGGCGGTCAGACACACCTGTTTATGCTGGGCG 956
DB 20887 GTGACAAAGCAGATGAACCCGCTATGAAAGAGGCTGTTCACATATGCCGTTGCC 20828
OY 957 AACATGCGCGCTGCTGCTGCGGAGAGCTGACCTACGCGCTTACCAACGCGCATGCCG 1016
DB 20827 AATATCCCTGGGCGGCTTCTGCGCACTTACACATGCGCCCTACCAATGCTCTCT 20768
OY 1017 TATGCTGCTGAGCTTGGCGGACCATGCT 1044
DB 20767 TATATGAAAGCTTGTGCTGCGCAAGAT 20740

RESULT 7
V74442/c
ID V74442: standard; DNA: 9280 BP.
AC V74442:
DE 16-MAR-1999 (first entry)
Staphylococcus aureus contig SEQ ID #131.
KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
KM toxic shock syndrome; ds.
OS Staphylococcus aureus.
FH Key Location/Qualifiers
FT misc-feature 841..900
FT /tag- a
FT /note- "these bases represent a line of missing text in
the sequence listing in the specification. They
are included to maintain the nucleotide numbering
given in the specification for this DNA sequence"
FT misc-feature 2641..2700
FT /tag- b
FT /note- "these bases represent a line of missing text in
the sequence listing in the specification. They
are included to maintain the nucleotide numbering
given in the specification for this DNA sequence"

```

```

FT misc-feature 4441..4500
FT /tag- C
FT /note- "these bases represent a line of missing text in
the sequence listing in the specification. They
are included to maintain the nucleotide numbering
given in the specification for this DNA sequence"
FT misc-feature 6241..6300
FT /tag- d
FT /note- "these bases represent a line of missing text in
the sequence listing in the specification. They
are included to maintain the nucleotide numbering
given in the specification for this DNA sequence"
FT misc-feature 8041..8100
FT /tag- e
FT /note- "these bases represent a line of missing text in
the sequence listing in the specification. They
are included to maintain the nucleotide numbering
given in the specification for this DNA sequence"

EP-786519-A2.
30-JUL-1997.
PR 07-JAN-1997; 100117.
PR 05-JAN-1996; US-009861.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA,
PI Rosen CA;
DR WPI: 97-374922/35.
PT Polynucleotide(s) and proteins derived from Staphylococcus aureus -
PT stored on computer readable medium and used in the production of
PT anti-S.aureus vaccines
PS Claim 1: Page 710-715; 3271pp; English.
CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the S.aureus DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or
CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against S.aureus infection. The
CC polypeptides can also be used in a kit for the immunodetection of
CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC skin and surgical wound infections, scalded skin syndrome, toxic shock
CC syndrome, etc. Organisms transformed with the DNA sequences can be used
CC for recombinant production of the polypeptides. The new DNA sequences
CC (and their fragments) are useful as primers or probes for isolating
CC homologues of any of the S.aureus DNA sequences contained on the
CC computer readable medium.
SQ Sequence 9280 BP; 2958 A; 1649 C; 1309 G; 3060 T;

Query Match 13.6%; Score 168; DB 1; Length 9280;
Best Local Similarity 50.6%; Pred. No. 1e-26;
Matches 484; Conservative 0; Mismatches 465; Indels 7; Gaps 3;

OY 151 CATGAGTGCATACCAAGCAGGTCGCGAGAGGCTGCTATACCGAGCGGATTC 210
DB 9280 CATACTGTTTATGTCGAACAATAATGCGGTTACAGATATCTTTGAAGATGATATAC 9221
OY 211 AAGCGGACGAGCGCGCACTGTCGGACCGCCGACCAAGTGTGGCGGACGCTATTA 270
DB 9220 AAAGAGAGCAGTGTGCGATATGCTGCTGTAACACGAAAGTTGG---GATGTGATATG 9164
OY 271 TTGCTCAAGTCAAGAACCGATAGCGGGAATAGCGCGCTGCGACAGCGGC-GATC 329
DB 9163 GTTATTAAGTTAAAGAACCACTGAATCTGAATATCAATTTTAAAGAGGCTGTA 1014
OY 330 TTGTCAGTCTCTGATTTGGCGCGTACAGCTGCTTCCACCATGCTGTTTGATTC 389
DB 9103 TTATTCATCTTATCTTCAATTTGCAAAATGAAGAAAATTAACCAAGCTTTGATGATGA 9044
OY 390 GGCACACGTAATGCTTACGAGACGCTCAGACCGCGCGGCGGACACTCCCTGCTT 449

```


Db 9043 AAGTAATTAGTATGATGACACTGTGCACTTACCAACCGACTTTCATTGTTA 8984
 QY 450 CCCCCATGCGCAAGTCCGCGGTCAGCTCGCCCGCAGGTTGGCTTACCACTGATG 509
 Db 8983 TCACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8924
 QY 510 CGAACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 569
 Db 8923 AACTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8864
 QY 570 GT 529
 Db 8863 GTAACTTTTTCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 8804
 QY 630 GGGCGACGCTTACGCTTACGCTTACGCTTACGCTTACGCTTACGCTTACGCTT 689
 Db 8803 GGTGCAATGTTTACGATTTTACGATTTTACGATTTTACGATTTTACGATTTT 8744
 QY 690 TTCTGCGGCGGATTCACACTGCTTACGCTTACGCTTACGCTTACGCTTACG 749
 Db 8743 TTCTGCGGCGGATTCACACTGCTTACGCTTACGCTTACGCTTACGCTTACG 8684
 QY 750 CTTGCGGCGGATTCACACTGCTTACGCTTACGCTTACGCTTACGCTTACGCTT 809
 Db 8683 CAAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8624
 QY 810 TCGAATTCATCTTGTGCGCATGTAACCAAGTGTGCTTACGCTTACGCTTACG 869
 Db 8623 ACAGAGACATGATTAACCAAGTGTGCTTACGCTTACGCTTACGCTTACGCTT 8564
 QY 870 CAGGCGGCTTTCGAAAGGCTC---ACGACCGACCTTACGCTTACGCTTACG 926
 Db 8563 CAAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8504
 QY 927 GTGACGACGCTTGTGCTTACGCTTACGCTTACGCTTACGCTTACGCTTACG 986
 Db 8503 AAGCATGTTGCTTACGCTTACGCTTACGCTTACGCTTACGCTTACGCTTACG 8444
 QY 987 ACCTACGCGCTGACCAACGCAAGTGTGCTTACGCTTACGCTTACGCTTACG 1046
 Db 8443 ACGTACGCTTAAATGATGATGATGATGATGATGATGATGATGATGATGATG 8384
 QY 1047 CGGGGCGGCTGCGATCCGCACTACGCAAGTGTGCTTACGCTTACGCTTACG 1102
 Db 8383 AGAGAGCATTTAAATCAATCAACATTCATTCATTCATTCATTCATTCATTC 8328

RESULT 8
 ID T67971 standard; DNA: 1074 BP.

AC T67971.
 DT 15-JUL-1997 (first entry)
 DE H. pylori membrane protein ORF 05cp20518orf61.
 KW Vaccines; prevention; treatment; infection; identification;
 KW binding compound; bacterium; life cycle; activator; bacteria;
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis;
 KW membrane; amino acid; metabolism; ds.
 OS Helicobacter pylori.
 FH Key Location/Qualifiers
 FT cds 1..1074
 FT /tag- a
 FT /note- "no stop codon given"
 PD MO9640893-A1.
 PD 19-DEC-1996: U09122.
 PF 06-JUN-1996: U09122.
 PR 07-JUN-1995: US-487032.
 PR 01-APR-1996: US-630405.
 PA (Aspr) AsprA AB.
 PI Berglindh OT, Smith D, Mellgaard BL.
 DR WPI: 97-052306/05.
 DR P-PDB: W20718.
 PT Helicobacter pylori nucleic acid sequences and related
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori

PT Infection, and to detect Helicobacter
 PS Claim 27; Page 819; 1481pp; English.
 CC The present sequence encodes a Helicobacter pylori membrane
 CC protein likely to contain four membrane spanning regions.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds,
 CC useful as potential H. pylori life cycle activators or inhibitors.
 CC The genomic sequence of H. pylori (ATCC 55679) was determined from
 CC overlapping contigs generated by mechanically shearing the
 CC bacterial DNA. The sequences were analysed for ORF of at least 180
 CC nucleotides, and the predicted coding regions defined by computer
 CC evaluation. To identify likely H. pylori antigens for vaccine
 CC development, the amino acid sequences predicted from various ORF
 CC were analysed for significant homology to other known or exported
 CC membrane proteins. Having identified and determined the sequences
 CC of interest, particular regions can be isolated from H. pylori by
 CC PCR amplification for recombinant polypeptide production, e.g. in
 CC E. coli hosts.
 SQ Sequence 1074 BP; 302 A; 189 C; 297 G; 286 T;

Query Match 7.1%; Score 87.6; DB 1; Length 1074;
 Best Local Similarity 46.4%; Pred. No. 4.2e-10;
 Matches 372; Conservative 0; Mismatches 414; Indels 16; Gaps 2;

QY 265 GATTATTGCTCAAGTCAAAAGACGATAGCGGCGGATAGCGGCGCTCGCA-CACGG 323
 Db 133 GATTATTGCTCAAAAGACGATAGCGGCGGATAGCGGCGCTCGCAAAAGAAA 192
 QY 324 GCGATCTGTTGCTGCTTGTGCTTGTGCGGCGGATAGCGGCGCTCGCAAAAGAAA 383
 Db 193 GCGACTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 252
 QY 384 GATTCCGCGACGATGATGCTTACGACGCGTCAAGCGCGGCGGCGGCGGCGGCGG 443
 Db 253 AATTAATAAATCACTTCTTATTTGCACTGAACCATTTGCGGCGCTTAAACACATACCT 312
 QY 444 CTGCTGCGCGGATGAGAC AGTCGCGGCTGACGCGGCGGCGGCGGCGGCGGCGG 495
 Db 313 ATTTCAGCGCTTATGAGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 372
 QY 495 -----GCTTACCACTGATGCGAACCAAGGCGGCGGCGGCGGCGGCGGCGGCGG 548
 Db 373 TTAAGCGCTTATGAGCGGCTTATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 432
 QY 549 CCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 608
 Db 433 TCGGCTGCGCAAGGCGGCTAATGCTGTAATGAGGCGGCGGCGGCGGCGGCGGCGG 492
 QY 609 GCCCGATGCGCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 668
 Db 493 GCGCAAGCTTTAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 552
 QY 669 CTTGCGCACTGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 728
 Db 553 TTACAAACCACTTATTCATTTGATGATTTGATTTGATTTGATTTGATTTGATTTG 612
 QY 729 GAGCTGAGGCTGCGTCAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 788
 Db 613 AATATCATTCAGCTTAAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 672
 QY 789 GCCAAGGCGCAATTAATGCTGCAATTCCTGCGGCGGCGGCGGCGGCGGCGGCGG 848
 Db 673 AGCCAAACCCCTAAAGTATGATTTAAGAGGCGATTTAAATACATGCGAGACAAAGGGA 732
 QY 849 CTGTTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 908
 Db 733 GTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 792
 QY 909 GACCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 968
 Db 793 TCTAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 852

Fri Jun 23 09:31:34 2000

us-09-362-485-6.rng

Page 12

Search completed: June 22, 2000, 15:09:11
Job time: 17877 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 15:08:43 ; Search time 341.15 Seconds
(without alignments)
470,560 Million cell updates/sec

Title: US-09-362-485-6
Perfect score: 1235
Sequence: 1 ATCTTGACAGTTATATGAC.....GGGAGACGATGATCGGCC 1235

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 6492525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued Patents_NA:*

- 1: /cgn2_6/ptodata/2/1na/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/2/1na/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/2/1na/5C.COMB.seq:*
- 4: /cgn2_6/ptodata/2/1na/5D.COMB.seq:*
- 5: /cgn2_6/ptodata/2/1na/6.COMB.seq:*
- 6: /cgn2_6/ptodata/2/1na/6C.COMB.seq:*
- 7: /cgn2_6/ptodata/2/1na/6ackfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	3.7	2582	4	US-08-816-105A-2
2	45.4	3.7	833	3	US-08-403-852D-3
3	45.4	3.7	5392	3	US-08-403-852D-1
4	43.6	3.5	78958	1	US-08-258-261B-6
5	43.6	3.5	28958	1	US-08-456-837-6
6	43.6	3.5	28958	1	US-08-457-342-6
7	43.6	3.5	28958	1	US-08-457-646A-6
8	43.6	3.5	28958	2	US-08-458-076A-6
9	43.6	3.5	28958	2	US-08-764-233A-4
10	43.6	3.5	28958	2	US-08-457-335A-6
11	43.6	3.5	28958	2	US-08-429-214-6
12	43.6	3.5	49377	2	US-08-764-233A-1
13	42.4	3.4	1833	3	US-08-403-852D-6
14	42.4	3.4	2185	1	US-08-173-508-3
15	42.4	3.4	2185	3	US-08-265-310-3
16	41.2	3.3	2414	7	5248599-1
17	41	3.3	1650	3	US-08-461-775-10
18	41	3.3	20235	2	US-07-642-734C-3
19	40.8	3.3	44377	4	US-08-804-227C-7
20	40.8	3.3	44377	4	US-08-804-198-1
21	40.6	3.3	4257	4	US-08-900-473-1
22	40.2	3.3	12001	2	US-08-458-568A-11
23	40.2	3.3	2064	1	US-08-343-428-1
24	40.2	3.3	11219	2	US-07-642-734C-1
25	39.6	3.2	474	3	US-08-403-852D-14
26	39.4	3.2	459	4	US-08-387-942C-35
27	39.4	3.2	1998	7	5212296-8

28	39.4	3.2	2658	3	US-08-461-775-11	Sequence 11, Appl
29	39.4	3.2	12588	4	US-08-387-942C-1	Sequence 1, Appl
30	39.2	3.2	420	1	US-08-470-179-148	Sequence 148, App
31	39.2	3.2	2353	6	PCT-US92-06840-1	Sequence 1, Appl
32	39.2	3.2	28804	3	US-08-592-874-1	Sequence 1, Appl
33	39	3.2	28804	5	US-09-096-942-2	Sequence 2, Appl
34	39	3.2	43280	3	US-08-804-227C-1	Sequence 1, Appl
35	38.8	3.1	8051	4	US-08-576-626A-2	Sequence 2, Appl
36	38.8	3.1	11219	2	US-07-642-734C-1	Sequence 1, Appl
37	38.6	3.1	1187	2	US-08-440-856A-2	Sequence 2, Appl
38	38.6	3.1	3231	2	US-08-074-121-4	Sequence 4, Appl
39	38.6	3.1	3231	6	PCT-US94-06447-4	Sequence 4, Appl
40	38.4	3.1	1215	5	US-08-947-726A-1	Sequence 1, Appl
41	38.4	3.1	2048	5	US-08-776-251-1	Sequence 1, Appl
42	38.2	3.1	2109	4	US-08-555-568B-20	Sequence 20, Appl
43	38	3.1	30001	4	US-08-125-468-1	Sequence 1, Appl
44	38	3.1	30001	3	US-08-474-933-1	Sequence 1, Appl
45	37.8	3.1	2588	3	US-08-796-414B-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-816-105A-2
Sequence 2, Application US/08816105A
Patent No. 5989882
GENERAL INFORMATION:
APPLICANT: Gruenger, Anneliese; Dellweg, Hans-Georg; Lenz,
APPLICANT: Jurgen; Schroeder, Werner; Pape, Hermann;
APPLICANT: Goekle, Klaus; Schaper, Beate; Hemker, Michael;
APPLICANT: Plepersberg, Wolfgang; Distler, Jurgen;
APPLICANT: Straetmann, Ansgar
TITLE OF INVENTION: PROCESSES FOR PREPARING ACARYOSYL
TITLE OF INVENTION: TRANSFERASE AND FOR USING IT IN THE
TITLE OF INVENTION: CONVERSION OF ACAROSE HOMOLOGUES
TITLE OF INVENTION: INFO ACAROSE, FOR THE PREPARATION
TITLE OF INVENTION: OF ACAROSE HOMOLOGUES
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Sprung Kramer Schaefer & Briscoe
STREET: 660 White Plains Road
CITY: Tarrytown
STATE: New York
COUNTRY: USA
ZIP: 10591-5144
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MB storage
COMPUTER: Apple Macintosh 6500
OPERATING SYSTEM: System 7.5
SOFTWARE: WordPerfect 3.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/816, 105A
FILING DATE: 14-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 19625269.5
FILING DATE: 25-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 19611252.4
FILING DATE: 22-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kurt G. Briscoe
REGISTRATION NUMBER: 33, 141
REFERENCE/DOCKET NUMBER: Bayer 9814-KGB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914) 332-1700
TELEFAX: (914) 332-1844
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2582 nucleotides
TYPE: nucleic acid
STRANDEDNESS: double

TOPOLOGY: linear
US-08-816-105A-2

Query Match 3.7%; Score 46; DB 4; Length 2582;
Best Local Similarity 47.3%; Pred. No. 0.013;
Matches 139; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

352 CCGGCTGACGCTCTGACACCGATGCGTTGTGATTCGGGACACGATCATGTCCTACG 411
1576 CCGGCTGCTGGGCGGCGGACCTTACCGGACGCTGCTGGGCGGCTGCGGCTGCGGAC 1635
412 AGACGCTCAGACCGGCGGCGGACGCTACCTGCTGCTGCGGATGACGGAAGTCGCGG 471
1636 TGACCGTGAGAGACGCGGACGCGGACGACGACGCTGCGGCGGCTGCGGCTGCGG 1695
472 GTGACGCTGCGGCGGCGGCTGCTGACGCTGATGCGAACCCAGGGGCGCGGCTG 531
1696 GCGAGGTCGCGCTGCTGCTGCTGACGCGGCGGCGGACGCGGCGGCTGCGGCGG 1755
532 TGCTGATGCGGCGGCGGCGGCGGCTGCGGACGCTGCGGATGCGGCGGCGG 591
1756 TGGGCGGCTGCTGACGCGGCGGCGGCGGACGCTGACGCTGCGGCGGCGGCTG 1815
592 CCGGCGGCTGACGACGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 645
1816 CCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1869

RESULT 2
US-08-403-852D-3
Sequence 3, Application US/08403852D
Patent No. 5891695

GENERAL INFORMATION:

APPLICANT: Blanc, Veronique
APPLICANT: Crouzet, Joel
APPLICANT: Jacques, Nathalie
APPLICANT: Lacroix, Patricia
APPLICANT: Thibaut, Denis
APPLICANT: Zagorec, Monique
APPLICANT: Debussche, Laurent
TITLE OF INVENTION: Polypeptides Involved In The
TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences
TITLE OF INVENTION: Coding For These Polypeptides And Their Use
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 10-MAY-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/FR 93/00923
FILING DATE: 25-SEP-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 92/11441
FILING DATE: 25-SEP-1992

ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.

REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03806, 0054-00000
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 833 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: S.pristineaspiralis

FEATURE:

NAME/KEY: CDS

LOCATION: 1..833

US-08-403-852D-3

Query Match 3.7%; Score 45.4; DB 3; Length 833;
Best Local Similarity 48.0%; Pred. No. 0.014;
Matches 130; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

542 CCGGCTGCGGCGGCGGCTGACGCGGCGGCGGCTGCTGATGCGGCGGCGGCGGCTA 601
327 CGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 386
602 CAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 661
387 CGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 446
662 CGACAACTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 721
447 ACTGACACGACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 506
722 GCGTACGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 781
507 CGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 566
782 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 812
567 CTTGCTGACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 597

RESULT 3

US-08-403-852D-1

Sequence 1, Application US/08403852D

Patent No. 5891695

GENERAL INFORMATION:

APPLICANT: Blanc, Veronique
APPLICANT: Crouzet, Joel
APPLICANT: Jacques, Nathalie
APPLICANT: Lacroix, Patricia
APPLICANT: Thibaut, Denis
APPLICANT: Zagorec, Monique
APPLICANT: Debussche, Laurent
TITLE OF INVENTION: Polypeptides Involved In The
TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences
TITLE OF INVENTION: Coding For These Polypeptides And Their Use
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS


```

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/403,852D
FILING DATE: 10-MAY-1995
PRIOR APPLICATION DATA: PCT/FR 93/00923
FILING DATE: 25-SEP-1993
APPLICATION NUMBER: FR 92/11441
FILING DATE: 25-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03806.0054-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5392 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: S.pristinaeapiralis
US-08-403-852D-1

```

```

Query Match          3.7%; Score 45.4; DB 3; Length 5392;
Best Local Similarity 48.0%; Pred. No. 0.023;
Matches 130; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 542 CCGGGTCCCGCGGCTGCAACCGCCGACGTCGTGATGCGCGCGGCAACCGCGGCTA 601
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3018 CGAGGGCGCGCGCGGCACTGACGTCGTCGCGCGCTGTGGGACGCTCGACGA 3077

QY 602 CAGCGAGCGCGGATGCGCAAGGCGGACCGGATACGGTTCTACATCAACAT 661
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3078 CGAGCGCTTGTCTACAGACCGCGCGGCTGTACTGCGCGGTGCCCCCGCTCACCA 3137

QY 662 CGACAACTTGGCAACTCGACGCGGAGTTCTGCGCGCGGATCCACACTCGTACTATC 721
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3138 ACTGACACCAAGGAGGAGCACTTCGACGTGGCGCGGCCCCCTCAAGCTGCGCGCGCC 3197

QY 722 GGCCTACGAGCTGAGAGGTCGCTCAAAAGTGCAGCTGCTGATTTGGGCGCTCTGT 781
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3198 GCAGGGCGCACCCGCTGTCGCGCTGACCGGCGCGGCTGCGCGCGCGCGACCTCGT 3257

QY 782 GCCAGGCGCAAGGACCCCAATTAGTCTCG 812
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3258 CTTGCTGACGAGGCGCGCGACGCGCGCTCG 3288

```

```

RESULT 4
US-08-258-261B-6
Sequence 6, Application US/08258261B
Patent No. 5639949
GENERAL INFORMATION:
APPLICANT: Schnupp, Thomas
APPLICANT: Ligon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Gaffney, Thomas Deane
APPLICANT: Lam, Stephen Ting
APPLICANT: Hammer, Phillip E.
APPLICANT: Uknes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
antipathogenic substances
NUMBER OF SEQUENCES: 22

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/258,261B
FILING DATE: 08-JUN-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-258-261B-6

```

```

Query Match          3.5%; Score 43.6; DB 1; Length 28958;
Best Local Similarity 44.8%; Pred. No. 0.098;
Matches 290; Conservative 0; Mismatches 354; Indels 4; Gaps 3;

QY 148 GGCATGAGGTGCTTCAATCCAGCACTGCGGAGAGGCGCTCGGCTTACCGAGCGGAT 207
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25306 GCCCAGAGAGACTACACCTCGCGGATGTGCTTATTCGTGCGCACACCGCGGCCAC 25365

QY 208 TTCAGCGCGGACGCGCGCACTGTGCGCACCGCGGACGAGTGTGGCGGACGCTGAT 267
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25366 TTCGAGCACCGCGCGCTCTGTGAGCCCAACCGGAGAGAGTCTCTCGGGGTGAC 25425

QY 268 TTATTTCTAAGTCAA-AGAACCGATAGCGGCGGGAATAGGCGCGCTCGGACAGGGCG 326
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25426 TCGCTGCGCCAGCAAGCCCGCGGACGACCGTCTCGGAGGAGGAGGAGCCAGCGGC 25485

QY 327 ATCTTGTACGTTCTTGATTTGGCGGCTGACGCTTGACCAAGTCTGTTGTGAT 386
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25486 AAGCTGCTCTGCTCTTCTTGCGCAAGGCTCGCACTGGGAAGGATGGCCCTCTCGTG 25545

QY 387 TCGGACCAAGCTCAATTGCTTACGAGACCGTTCAGACCGCGGACGCGCACTACCCCTG 446
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25546 CTGACTCTCGCGCTCTCTGCGCTGAGTGAAGCATGAGGAGCGCGCTCGCTCTCT 25605

QY 447 CTGCGCCGATGAGCAAGTTCGCGGCTGACGACGCGCGCGGCGGCTTGGGCTTACCAAC--C 504
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25606 CAGTGTGATGAGCTGCTGCTGCTGCGGCAAGGCTCGCACTGGGAAGGATGGCCCTCTCGAG 25665

QY 505 TGATGGAACCCAGGCGGCGGCGGTGCTGATGCGCGGCGGCTGCGGCGCTGCAACCGG 564
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25666 CGGCTGAGCTGTACAGCCCGGCGCTTGGCGTATGAGTGTCTCCCTGGGCGGCTCTGG 25725

QY 565 CCGAGC-TGCTGTATGCGCGCGCGGCGGCGGCTTACAGGACGACCGCATCGCAAC 623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25726 CGCTGCTGCGGTAGAGCCCGCGGCTGTGTCGCGCACAGTCAGGGGCGAGATCGCGGC 25785

```

QY	634	GGCAATGGGGCGACCGGTATCGGTTGTAGCAATCAACATGACAAACTTGGCGAACTCGAC	683
Db	25786	GGCTTCGTGCGAGGGGCTCTCTCCCTCGAGAGACGGGGCCGCATGCGCCCTCGCGACAC	25845
OY	684	GCCGAGTTCTGGGGCGGATGCACACTGCTACTATGGGCGCTTAGAGCTGAGGGTGGC	743
Db	25846	AAAGGCTTACCAACGTCGGCGGCAAGGGGGCGATGAGCGCGCTGTGAGCTCGGGCGCTTC	25905
OY	744	GTCAAACGTGCGACCTGTGATTTGGGGCGCTTCGTGTCACAGGGGCC	791
Db	25906	GACCTTCGACACTTACTCTGCTCTCTGGGGGCGAAGGCTTCCATCGCC	25953

RESULT 5

US-08-436-837-6
Sequence 6, Application US/08456837
Patent No. 5643774
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Gaffney, Thomas Deane
APPLICANT: Lam, Stephen Ting
APPLICANT: Hammer, Phillip E.
APPLICANT: Uknes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
NUMBER OF INVENTION: antipathogenic substances
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:

ADDRESSSEE: Ciba-Geigy Corporation
 STREET: 7 Skyline Drive
 CITY: Hawthorne
 STATE: NY
 COUNTRY: USA
 ZIP: 10532
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent Release #1.0, Version #1.252
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/456,837
 FILING DATE: 01-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/457,205
 FILING DATE: 01-JUN-1995
 APPLICATION NUMBER: 08/258,261
 FILING DATE: 08-JUN-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Elmer, James Scott
 REGISTRATION NUMBER: 36,129
 REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8614
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 28958 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHEetical: NO
 ANTI-SENSE: NO
 US-08-456-837-6

Query Match	3.5%;	Score 43.6;	DB 1;	Length 28958;
Best Local Similarity	44.8%;	Pred. No. 0.098;		

	Matches	290;	Conservative	0;	Mismatches	354;	Indels	4;	Gaps	3;
OY	148	GGCCATGAGGGTCTCATCCAGGCAGAGTGGCCGAGAGGGCTCGGGTATATACCCAGCCGAT	207							
Db	25306	GCCCAAGACACTACCTTCGGGATGGGCTTATTCGGTGGCCACACACCCGGCCAC	25365							
OY	208	TTCAAGGGCGGACGGCGCGCACTGGTCGGCACCGCCGACACAGTGTGGGCGCACGCTGAT	267							
Db	25366	TTTCAGACACCGGGCGCTCTCTGATCCCAACACCGGACGACACTCTCTCGCGGCTCAG	25425							
OY	268	TTATTGCTCAGAGTAA-AGAACCATATGCGGGCAATATCGCCGCTTGCGACACGGGCG	326							
Db	25426	TCGCTGGCCCAAGACAACGCCGCCCGACACCGCTCTCGACGAGGAGGAAAGCCACGGC	25485							
OY	327	ATCTGTTCACGCTTTTGATTTGGCCGAGTCACTGTCTTGACCGATGCGTTGTGGAT	386							
Db	25486	AAGCTCGTCTTGCTTTTCTTGCGCAAGGCTTCGACGTGGGAAGGATGGCCCTTCGCTG	25545							
OY	387	TCGCGACACAGCTCAATTGCTACGAGACCGTCCAGACCGCCGACGGCGCACTACCCCTG	446							
Db	25546	CTCGACTCTCCGCCGCTCTTCGCCCTCAGCTCGAAGCATGGACGACGGCGCTCGCTCT	25605							
OY	447	CTTGCCCGGATAGGGAAGTGGCGGGTGGACTCGCGGCCAGTGGGGGCTATACAC--C	504							
Db	25606	CACGTCGAGTAGGCTGCTGCGCTCTCTGCGCCGAGAGAGGGGCCCTCTCCCTCGAC	25665							
OY	505	TGATCGAAACCCAAAGGGGCGCGGCTGTGCTGATGGCGGGGTGCCCGCGCTCGAAACCG	564							
Db	25666	CGCGTCGACGTGTACACGCCGCCCTCTTTGCGGTATGATGTCTCCCTGGCGCCCTCTGG	25725							
OY	565	CCGAGC-TCGTGATATGGCGCGCGCGCACCGCGGCTTACACGACGCCGCTATGCCCAAC	623							
Db	25726	CGCTGCTGGCGGTAGAACGCCGCCGCCGCTGTGCGGACACAGTCAGGGGAGATGTCGCCGC	25785							
OY	624	GGCATGGGCGGCACGTTACGCTTTAAGACATCAATCGACAAACTTGTGGCAATCTGCAC	683							
Db	25786	GCCTTCTGTGCGAGGGGCTCTCTCCCTTCGAGAGCGGGCGCCGATGCGCCGCTTGCGCAC	25845							
OY	684	GCCGAGTTCTGGGCGCGGATCCACACTGCTACTCATCGGCTTACGAGCTGAGGGGTGCC	743							
Db	25846	AAAGGCTCACACACGTCGCGCGGCAACGAGGCGCATATGCGCGCCGTGTGAGCTCGGCGCTCC	25905							
OY	744	GTCAAACGTGGCGAGCTGGTGATTGGGGCGGTCGTGATGTCACAGGGGCC	791							
Db	25906	GACCTTCAGACTACTCTGCTCTGTGGGGGACACAGGCTCTTCATATGCC	25953							

RESULT

```

US-08-457-342-6
: Sequence 6, Application US/08457342
: Patent No. 5662898
:
: GENERAL INFORMATION:
:
: APPLICANT: Schupp, Thomas
:
: APPLICANT: Ligon, James M.
:
: APPLICANT: Beck, James Joseph
:
: APPLICANT: Hill, Dwight Steven
:
: APPLICANT: Ryals, John Andrew
:
: APPLICANT: Gaffney, Thomas Deane
:
: APPLICANT: Lam, Stephen Ting
:
: APPLICANT: Hammer, Phillip E.
:
: APPLICANT: Urnes, Scott Joseph
:
: TITLE OF INVENTION: Genes for the synthesis of
:
: TITLE OF INVENTION: antipathogenic substances
:
: NUMBER OF SEQUENCES: 22
:
: CORRESPONDENCE ADDRESS:
:
: ADDRESSEE: Ciba-Geigy Corporation
:
: STREET: 7 Skyline Drive
:
: CITY: Hawthorne
:
: STATE: NY
:
: COUNTRY: USA
:
: ZIP: 10532
:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/457,342
 FILING DATE: 01-JUN-1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/457,205
 FILING DATE: 01-JUN-1995
 APPLICATION NUMBER: 08/258,261
 FILING DATE: 08-Jun-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Elmer, James Scott
 REGISTRATION NUMBER: 36,129
 REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8614
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 28958 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-457-342-6

Query Match 3.5%; Score 43.6; DB 1; Length 28958;
 Best Local Similarity 44.8%; Pred. No. 0.098;
 Matches 290; Conservative 0; Mismatches 354; Indels 4; Gaps 3;

148 GCCCATGAGTCTCATCAGGAGGCGGAGAGGGCTCGCTATCAGCCAGCGGAT 207
 25306 GCCCAGCGACCTTCACCTCGCGATGTGGCTATTGCTGCTACCCAGCCGGCCAC 25365
 208 TTCAAGGGGAGGCGGCACTGCTGGGCGACCGCCGACAGTGTGGCGGAGCTGAT 267
 25366 TTGAGCAGCCCGCGCTCTGTAGCCACACACCGGACGACGCTCTCGGCTGAC 25425
 268 TTATGCTCAAGTCAA-AGAACGATAGCGGGGAATAGCGCCCTCGGACACGGGCG 326
 25426 TCGTCGCGGAGGAGAACCGCCCGCCGACCGCTCTCGGAGCGAGCGAGCGGCG 25485
 327 ATCTTTGTCAGCTTCTGCAATTTGGCCCGCCGTCAGTGTGACCGATGGCTTGAT 386
 25486 AAGTCGCTTCTGCTTCTGCGGAGGCTGCGAGTGGAGGAGTGGCCCTTCGCTG 25545
 387 TCCGGACACGATCAATGCTAGAGAGACCGTCCAGACCGCGGCGGCACTACCCCTG 446
 25546 CTCGACTCTGCGCGCTTTCGGGCTCAGCTCGAAGCATCGAGCGGCGGCTC 25605
 447 CTTCGCGGATGAGGAGTCCGCGGTGACTCGCCGCCGAGTTGGCGCTTACAC--C 504
 25606 CACGTCGAGTGGAGCTGCTGCTGCGCGCGAGCA GCGCGCTCTCTCTCTGAC 25665
 505 TGATCGGAACCAAGGGGCGCGGTGCTGATGGGCGGGGTCCCGGCTCGAACCGG 564
 25666 CGCGTCGACGTCGTACACACCGCCCTTTTGGCGTCAAGGTCCTCGGCGGCTCTGG 25725
 565 CCGAGC-TCGTGATGATGCGCGCGGACCGCGGCTACACAGACGACCGCATGCCAAC 623
 25726 CGCTGCTGCGGTAGAGACCGCGCGCTGTGCGGCCACATCAGGGGAGATGCGGCC 25785
 624 GGCATGGGCGGAGCGTACGTTTACATCAACATCGACCAAACTTGGCAACTGAC 683
 25786 GCTTCGTGCGAGGCGCTCTCTCTCGAGAGACGCGCGCCGATGCGGCCCTGCGCAGC 25845
 684 GCCGATCTGCGGCGGATCCACACTGCTACTGATGCGGCTTACGAC CGAGGGTGGC 743

Db 25846 AAAGGCTACACCGTGGCGGCAAGGGGCGATGGCCCGCTCGAGCTCGGCGCTCC 25905
 Oy 744 GTCAAAGCTGCCAGCTGTGATGGGCGCGCTCTGTGCGCAGCGCC 791
 Db 25906 GACCTCCAGACTTACTGCTTCCCTGGGCGGACAGGCTCTTCATCGCC 25953

RESULT 7
 US-08-457-646A-6
 Sequence 6, Application US/08457645A
 Patent No. 5679560
 GENERAL INFORMATION:
 APPLICANT: Schnupp, Thomas
 APPLICANT: Ligon, James M.
 APPLICANT: Beck, James Joseph
 APPLICANT: Hill, Dwight Steven
 APPLICANT: Ryals, John Andrew
 APPLICANT: Gaffney, Thomas Deane
 APPLICANT: Lam, Stephen Ting
 APPLICANT: Hammer, Phillip E.
 APPLICANT: Uknes, Scott Joseph
 TITLE OF INVENTION: Genes for the synthesis of
 TITLE OF INVENTION: antipathogenic substances
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Ciba-Geigy Corporation
 STREET: 7 Skyline Drive
 CITY: Hawthorne
 STATE: NY
 COUNTRY: USA
 ZIP: 10532
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-/S/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/457,646A
 FILING DATE: 01-JUN-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/457,205
 FILING DATE: 01-JUN-1995
 APPLICATION NUMBER: 08/258,261
 FILING DATE: 08-Jun-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Elmer, James Scott
 REGISTRATION NUMBER: 36,129
 REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8614
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 28958 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-457-646A-6

Query Match 3.5%; Score 43.6; DB 1; Length 28958;
 Best Local Similarity 44.8%; Pred. No. 0.098;
 Matches 290; Conservative 0; Mismatches 354; Indels 4; Gaps 3;

148 GCCCATGAGTCTCATCAGGAGGCGGAGAGGGCTCGCTATCAGCCAGCGGAT 207
 25306 GCCCAGCAGACTTCACCTTCGCGATGTGGCTATTGCTGCGCACACCGCGGCCAC 25365
 208 TTCAAGGGGAGGCGGCACTGCTGGGCGACCGCCGACGATGTGGCGGAGCTGAT 267

```

DB 25366 TTCGAGACCGGCGGCTCTGAGCCCAACCGGAGAGACTCTCTCCGCGCTCGAC 25425
OY 268 TTATGCTCAAGGTCAA-AAGAACGATAGCGCGGGAATAGCGCCCTCGACACGGCGG 326
DB 25426 TCGCTCGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 25485
OY 327 ATCTGTTACAGTCTTGTGATTTGGCCGGCTGACGCTGCTGACCGATGCTGTTGGAT 386
DB 25486 AAGCTGCTGCTGCTCTTCTGAGCAAGGCTGCGAGTGGGAAGAGTGGCCCTCGGCTG 25545
OY 387 TCCGACACCGCTCAATGCTGCTAGAGACCGCTCAAGACCGCGGCGCACTACCCCTG 446
DB 25546 CTCGACTCTCTCCGCTCTCCGCGCTGAGCTCGAAGCATGAGAGCGCGCTCGCTCT 25605
OY 447 CTGCGCCGATAGAGCAAGTCCCGGCTGAGCTCGCGCCCAAGTGGCGCTTACCAAC--C 504
DB 25606 CACGTCGAGTGGAGCGCTCCCGCTGCTGCGCCGAGAGAGAGAGAGAGAGAGAGAG 25665
OY 505 TGATGGAGAACCAAGGAGGCGGCTGCTGATGGGCGGG-TGGCCGCGCTGCAACCGG 564
DB 25666 CCGCTGAGCTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 25725
OY 565 CCGAGC-TCGTGCTGATCGGCGGCGGCGGCGGCTACACGAGAGAGAGAGAGAGAGAG 623
DB 25726 CGCTGCGCTGCGGCTGAGAGCGCGCGGCGGCTGCTGCGGCAAGTACAGAGAGAGAG 25785
OY 624 GCGATGAGGCGGCGGCTGAGGCTTCTAGACATCAACATGCAACAACTTGGCACTCGAC 683
DB 25786 GCGTCTGCTGAGGCGGCTGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 25845
OY 684 GCGGAGTTCGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 743
DB 25846 AAGCGCTACACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25905
OY 744 GTCAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 791
DB 25906 GACCTTCAGACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25953

```

RESULT 8

US-08-458-076A-6

Sequence 6, Application US/08458076A

Patent No. 5698425

GENERAL INFORMATION:

APPLICANT: Schupp, Thomas

APPLICANT: Lyon, James M.

APPLICANT: Beck, James Joseph

APPLICANT: Hill, Dwight Steven

APPLICANT: Ryals, John Andrew

APPLICANT: Gaffney, Thomas Deane

APPLICANT: Lam, Stephen Ting

APPLICANT: Hammer, Phillip E.

APPLICANT: Unnes, Scott Joseph

TITLE OF INVENTION: Genes for the synthesis of

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

STREET: 7 Skyline Drive

CITY: Hawthorne

STATE: NY

COUNTRY: USA

ZIP: 10532

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08-458-076A

FILING DATE: 01-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/457, 205

FILING DATE: 01-JUN-1995

APPLICATION NUMBER: 08/258, 261

FILING DATE: 08-JUN-1994

ATTORNEY/AGENT INFORMATION:

NAME: Elmer, James Scott

REGISTRATION NUMBER: 36,129

REFERENCE/DOCKET NUMBER: CGC 1506/CIP3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8614

TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 28958 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHEICAL: NO

ANTI-SENSE: NO

US-08-458-076A-6

Query Match 3.5% Score 43.6; DB 2; Length 28958;

Best Local Similarity 44.8%; Pred. NO. 0.098;

Matches 290; Conservative 0; Mismatches 354; Indels 4; Gaps 3;

```

OY 148 GCGCATGAGTGTCTATCAGGAGAGTGGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 207
DB 25306 GCGCATGAGTGTCTATCAGGAGAGTGGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 25365
OY 208 TTCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 267
DB 25366 TTCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 25425
OY 268 TTATGCTCAAGGTCAA-AAGAACGATAGCGCGGGAATAGCGCCCTGCGACAGGCGG 326
DB 25426 TCGCTCGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 25485
OY 327 ATCTGTTACAGTCTTGTGATTTGGCCGGCTGACGCTGCTGACCGATGCTGTTGGAT 386
DB 25486 AAGCTGCTGCTGCTCTTCTGAGCAAGGCTGCGAGTGGGAAGAGTGGCCCTCGGCTG 25545
OY 387 TCCGACACCGCTCAATGCTGCTAGAGACCGCTCAAGACCGCGGCGCACTACCCCTG 446
DB 25546 CTCGACTCTCTCCGCTCTCCGCGCTGAGCTCGAAGCATGAGAGAGAGAGAGAGAGAG 25605
OY 447 CTGCGCCGATAGAGCAAGTCCCGGCTGAGCTCGCGCCCAAGTGGCGCTTACCAAC--C 504
DB 25606 CACGTCGAGTGGAGCGCTCCCGCTGCTGCGCCGAGAGAGAGAGAGAGAGAGAGAGAG 25665
OY 505 TGATGGAGAACCAAGGAGGCGGCTGCTGATGGGCGGCTGCGGCGCTGCAACCGG 564
DB 25666 CCGCTGAGCTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 25725
OY 565 CCGAGC-TCGTGCTGATCGGCGGCGGCGGCGGCGGCTACACGAGAGAGAGAGAGAGAG 623
DB 25726 CGCTGCGCTGCGGCTGAGAGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25785
OY 624 GCGATGAGGCGGCGGCTGAGGCTTCTAGACATCAACATGCAACAACTTGGCACTCGAC 683
DB 25786 GCGTCTGCTGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25845
OY 684 GCGGAGTTCGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 743
DB 25846 AAGCGCTACACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25905
OY 744 GTCAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 791
DB 25906 GACCTTCAGACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25953

```

```

1      RESULT 9
2      US-08-764-233A-4
3      Sequence 4, Application US/08764233A
4      Patent No. 5716849
5      GENERAL INFORMATION:
6      APPLICANT: Ligon, James M.
7      APPLICANT: Schupp, Thomas
8      APPLICANT: Beck, James J.
9      APPLICANT: Hill, Dwight S.
10     APPLICANT: Neff, Snezana
11     APPLICANT: Ryals, John A.
12     TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen
13     NUMBER OF SEQUENCES: 10
14     CORRESPONDENCE ADDRESS:
15     ADDRESSEE: Ciba-geigy Corporation
16     STREET: 520 White Plains Road, P.O. Box 2005
17     CITY: Tarrytown
18     STATE: NY
19     COUNTRY: USA
20     ZIP: 10591
21     COMPUTER READABLE FORM:
22     MEDIUM TYPE: Floppy disk
23     COMPUTER: IBM PC compatible
24     OPERATING SYSTEM: PC-DOS/MS-DOS
25     SOFTWARE: Patentin Release #1.0, Version #1.30
26     CURRENT APPLICATION DATA:
27     APPLICATION NUMBER: US/08/764,233A
28     FILING DATE:
29     CLASSIFICATION: 435
30     PRIOR APPLICATION DATA:
31     APPLICATION NUMBER: US 08/729,214
32     FILING DATE: 09-OCT-1996
33     PRIOR APPLICATION DATA:
34     APPLICATION NUMBER: US 08/258,261
35     FILING DATE: 08-JUN-1994
36     ATTORNEY/AGENT INFORMATION:
37     NAME: Weigs, J. Timothy
38     REGISTRATION NUMBER: 38, 241
39     REFERENCE/DOCKET NUMBER: 1506/CIP6
40     TELECOMMUNICATION INFORMATION:
41     TELEPHONE: (919) 541-8587
42     TELEFAX: (919) 541-8689
43     INFORMATION FOR SEQ ID NO: 4:
44     SEQUENCE CHARACTERISTICS:
45     LENGTH: 28958 base pairs
46     TYPE: nucleic acid
47     STRANDEDNESS: single
48     TOPOLOGY: linear
49     MOLECULE TYPE: DNA (genomic)
50     ORIGINAL SOURCE:
51     ORGANISM: Sorangium cellulosum
52     IMMEDIATE SOURCE:
53     CLONE: P98/1
54     US-08-764-233A-4

```

	Query Match	Similarity	3.5%	Score	43.6	DB 2:	Length	28958;
	Best Local	Similarity	44.8%	Pred. No.	0.098			
	Matches	290;	Conservative	0;	Mismatches	351;	Indels	4; G _A T _S
OY	148	GGCATGAGTGTCTCATCCAGCAGGTGCCGGAGAGGCGTCGCGTATCAACCGACGGGAT	207					
Db	23306	GCCCACGACAGACTCTACCCTCGCGGATGTGGCTATTGCTGGCCACACACCCCGCCCCAC	25365					
OY	208	TTCAAAGGGCGCAGGGCGGCACTGCTGGGACCGCCGACAGGTGTGGGCGCACGCTGAT	267					
Db	25366	TTCCGACACCGCGCGCTTTCGTAGCCACAACCGCAGAGAGTCTCTTCGCGGCTCGAC	25425					
OY	268	TTATTGCTCAAGGTCAA-AGAACCGATTACCGCGGGAATACGGCGCGCTCGACACCGGGG	326					
Db	25426	TGGCTGTGCCAAGACAAAGCCGCGCCGACACACGTCCTCTGGAGAGGAGCGGAACCCACGGG	25485					
OY	327	ATCTTGTTACGCTTCTTGCAATTTGGCCGCGTCACGTGCTT-ACCGATGCGTGTGGAT	386					

[illegible]

RESULT 10
 US-08-457-335A-6
 Sequence 6, Application US/08457335A
 Patent No. 5723759
 GENERAL INFORMATION:
 APPLICANT: Schnupp, Thomas
 APPLICANT: Ligon, James M.
 APPLICANT: Beck, James Joseph
 APPLICANT: Hill, Dwight Steven
 APPLICANT: Ryals, John Andrew
 APPLICANT: Gaffney, Thomas Deane
 APPLICANT: Lam, Stephen Ring
 APPLICANT: Hammer, Phillip E.
 APPLICANT: Unnes, Scott Joseph
 TITLE OF INVENTION: Genes for the synthesis of
 TITLE OF INVENTION: antipathogenic substances
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Ciba-Geigy Corporation
 STREET: 7 Skyline Drive
 CITY: Hawthorne
 STATE: NY
 COUNTRY: USA
 ZIP: 10532
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/457,335A
 FILING DATE: 01-JUN-1995
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/457,205
 FILING DATE: 01-JUN-1995
 APPLICATION NUMBER: 08/258,261
 FILING DATE: 08-Jun-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Elmer, James S.,att
 REGISTRATION NUMBER: 36,129

```

1 REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
2 TELECOMMUNICATION INFORMATION:
3 TELEPHONE: 919-541-8614
4
5 TELEREX: 919-541-8689
6
7 INFORMATION FOR SEQ ID NO: 6:
8
9 SEQUENCE CHARACTERISTICS:
10
11 LENGTH: 28958 base pairs
12
13 TYPE: nucleic acid
14
15 STRANDEDNESS: single
16
17 TOPOLOGY: linear
18
19 MOLECULE TYPE: DNA (genomic)
20
21 HYPOTHETICAL: NO
22
23 ANTI-SENSE: NO
24
25 US-08-457-335A-6

```

Query Match	score 43.6;	DB 2;	length 28958;
Best Local Similarity	44.8%;	Pred. No. 0.098;	
Matches 290; conservative	0;	Mismatches 35;	Indels 4; Gaps 3

Oy	148	GGCATTAGAGTCTCTATCCAGGACAGGTGGCGGAGAGGGCTCGTACTACCCAGCGGAT	207
Db	25306	GCCCAAGAGACTACACCTCGCGGATGAGGCTATTGGCTATGGCCACACCGCGCGCCAC	25365
Oy	208	TTCAAGGCGGACAGGCGCGCACTGGTCGGACACCGCGGACACAGTGTGGGCGCACCTGAT	267
Db	25366	TTTCGAGACCGGCGCGCTCTCGTACCCCAACACCGGAGAGACTCTCTCTCCGCGCTGCAC	25425
Oy	268	TTATTGCTCAAGCTCAA-AGAACCAGATAGCGGCGGAAATACGCGCCGCTTCGACAGGGCG	326
Db	25428	TGCGTCGCCCAAGAGACGGCCCCCGGACACCGCTCTCGAGCGGAGCGGAAGCCACGGC	25485
Oy	327	ATCTGTTCACCTTCTTGATTTTGGCGCGGTACAGTCTTGACCGATGCCATTGTTGGAT	386
Db	25486	AAGCTGTATTGCTCTTCTCTGGGCAAGGCTTCGAGTGGG-AGGATAGCCCTTCTCGTG	25545
Oy	387	TCCGGACACACGTCAATTGCTTACGAGACCGTCCAGACCGCGGACGCGCATACCCCTG	446
Db	25546	CTCGACTCTCCCGCTCTTCCGCGCTCAGCTCGAAGCATGCGACGGCGCTCCCTCT	25605
Oy	447	CTTGGCCCCATGAGCAATTCGCGCGGTGCACATCGCGCGCCAGGTGGGCTTACACAC--C	504
Db	25606	CACGTGAGATGAGACTGCTCGCGCTCCTCGCGCGGAGAGAGGGGCGCCCTCTCCCTGCAC	25665
Oy	505	TCATCGCAACCAAGGAGGCGCGGGTGTGCTGATGGGCGGGTGCCCGCGCTGCACCGG	564
Db	25666	CCGCTGAGAGTGTACAGCGCCGCCCTTGTGCGTATGTTCTCCCTGGGCGCCCTGTGG	25725
Oy	565	CCGACG-TCGTGTGATTCGCGCGCGGACCGCGCGGTACACAGCACGCCGCAATCGCCAAC	623
Db	25726	CGCTGCTGTGGGATAGAGCGCGCGCGCTGCTGGGCGCACAGTCAAGGCGGAGATCGCGCG	25785
Oy	624	GGCATGGGCGGACCTTACGGTTCTAGACATCAACATGACACAACTTGGGCACTCGAC	683
Db	25786	GGCTTGTGTGAGAGCCCTCTCTCCCTCGAGGAGCGCGCCGATCGCGCCCTGGCCACG	25845
Oy	684	GCCGATTTCTGGCGCGGATTCACACTCCCTACTCATCGGCTACGAGTCAGAGGATGGC	743
Db	25846	AAAGGCTTACACACGTCGCGGCAACGGGAGCCATGGCCGCGTGAAGCTCGGCGGCTTC	25905
Oy	744	GTCAAACGTGCGGACTGTGTGATTTGGGGCGGCTCTGTGTGCAAGCGCC	791
Db	25906	GACCTTCAAGCTACCTGCTCTCCCTTGGGGAGAGGCTCTTCATATGCC	25953

RESULT 11
US-08-729-214-6
Sequence 6, Application US/08729224
Patent No. 5817502
GENERAL INFORMATION:
APPLICANT: Ligon, James M.
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

Query Match	3.5%	Score 43.6;	DB 2;	length 28958;
Best Local Similarity	44.8%;	-Pred. No. 0.098;		
Matches 290;	Conservative	0;	Mismatches 354;	Indels 4;
				Gaps 3;

OY	148	GGCCATGAGTGGCTCTATCCAGGCAAGCCCGAGAGGAGGATCGGCTATACACCGAGGAT	207
Db	25306	GCCACAGCAGCACTTACCCTCCGGATGTGGCTTATTCGTGGCCACACACCGCGCCAC	25365
OY	208	TTCAAGCGGAGCAGCCCGCACCTGGTTCGGACCGCGCCAGCCAGGTGTGTGGCCGACGCTGAT	267
Db	25366	TTTGAGACCGGGCGCTCTGTAGCCACAAACGGCGACAGCTCTCTCCGGGCTCGAC	25425
OY	268	TTATTTGCTCAAGTCAA -AGAACCATATGCGCGGGAATAGCGCCGCTCGACACAGGCG	326
Db	25426	TGCGTCCCCGAGACAAAGCCGCCCGCGAGCACCGCTCTCGGACGAGACGAAACCCAGCG	25485
OY	327	ATCTTTTCACGTTCTTGCAATTTGGCCCGATCAGTGGTTCTACACCATGGCTTTGGAT	386
Db	25486	AACTGCTGCTTTCGCTCTTCTCTGGCCAAAGCTGGACAGTGGAAAGGATGGCCCTCTGTGCTG	25545
OY	387	TCCGCGACACAGTCATATTCGTACAGAGACGTCCAGACCGCGCGACATCACTCCG	446
Db	25546	CTGACATCTGTGGCCGCTCTTCGGCGCTCAGGTCGAAGATATCGAGCGCGGCTCGCTCTCT	25605
OY	447	CTTCCCCCGATGAGCCAAAGTCGCGCGGTGACTCGCGGCCACAGGTTGGCGCTTACAC -C	504
Db	25606	CACGTCAAGTGGAGCCCTGGCTCGCGCTCTGCGCCGCGACGAGGAGGCGCCCTCTCCCTGAC	25665
OY	505	TGATGGACAAACCAAGGGGCGGGGGTGTGTATATGGCGGGGTCCCGGCGTGGAAACCG	564
Db	25666	CGGCTCTACGATCAAGCCCGCCCTTTGGCGCTATGATGTCTCCCTGGGAGCGCCCTCTGG	25725

```

0Y      565  CCGAGC-TGCTGTATCGGCGCGGACACCGCGGCTACAAAGCAGCCGACATCGCCAC 623
Db 25726 CGCTGCTCGGCGGTAGAGACCCGCCGCCCTGCTGCGGCACAGTCAGAGGCGAGATCGCCGC 25785
0Y      624  GGCATGGCGCGACCGTTACGGTTCTAGACATCAACATCGACAAACTTCGGCACTCGAC 683
Db 25786 GCGTTCGTGCGAGGCGCTCTCCCTCGAGACGGGCGCGCATGCGCCCTGCGGAC 25845
0Y      684  GCGGATCTCTGGGCGCGGATCCACACTGCTCTCATGCGGCTTAGAGGTGAGGCTGCG 743
Db 25846 AAGGCGCTACACCGTCGCGCGGCAACGGGCGCATGCGCGCGTGAAGTCTGCGCGCTTC 25905
0Y      744  GTCAAACGTGCGCGACCTGTGATGTGGGCGCGCTCTGTGTCACAGCGGCC 791
Db 25906 GACCTTCAGACCTACCTCGCTCCGAGGGGCGACAGGCTCTCATCGCC 25953

RESULT 12
US-08-764-233A-1
: Sequence 1, Application US/08764233A
: Patent No. 5716849
:
GENERAL INFORMATION:
APPLICANT: Ligon, James M.
APPLICANT: Schupp, Thomas
APPLICANT: Beck, James J.
APPLICANT: Hill, Dwight S.
APPLICANT: Neff, Snezana
APPLICANT: Ryals, John A.
TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591
:
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/764,233A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/729,214
FILING DATE: 09-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: 1506/CIP6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8687
TELEFAX: (919) 541-8689
:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4937 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Sorangium cellulosum
IMMEDIATE SOURCE:
CLONE: p98/1, pUL3, and pVKM15
FEATURE:
NAME/KEY: misc_feature
LOCATION: 383..760

```

```

OTHER INFORMATION: /product= "SorB"
OTHER INFORMATION: /note= "This gene encodes a protein that is highly homologous to type I PKs such as eryA from Saccharopolyspora erythraea."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 927..19874
OTHER INFORMATION: /product= "Sora"
OTHER INFORMATION: /note= "Gene product is highly homologous to type I PKs t are known to be involved in the synthesis of polypeptide compounds."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 942..7115
OTHER INFORMATION: /product= "Module 1 of Sora"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7203..12884
OTHER INFORMATION: /product= "Module 2 of Sora"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 13455..19615
OTHER INFORMATION: /product= "Module 3 of Sora"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19871..46318
OTHER INFORMATION: /product= "SorB"
OTHER INFORMATION: /note= "Gene product is highly homologous to type I PKs ge"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19870..24556
OTHER INFORMATION: /product= "Module 1 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 24638..30820
OTHER INFORMATION: /product= "Module 2 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 30881..35446
OTHER INFORMATION: /product= "Module 3 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 35328..40114
OTHER INFORMATION: /product= "Module 4 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 40190..46318
OTHER INFORMATION: /product= "Module 5 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 46851..47891
OTHER INFORMATION: /product= "SorM"
OTHER INFORMATION: /note= "The protein encoded by the sorM gene is highly homologous to the methyltransferase from Streptomyces hygroscopicus that is involved in the synthesis of the polyketide rapamycin."
US-08-764-233A-1

Query Match      3.5%; Score 43.6; DB 2: Length 49377;
Best Local Similarity 44.8%; Pred. No. 0.11; 354; Indels 4; Gaps 3;
Matches 290; Conservative 0; Mismatches 354; Indels 4; Gaps 3;

OY   148 GCCCATGAGTGTCTATCATCAGCAGAGTGCCGAGAGGGCTCGCTATCACCGACGCCGAT 207
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    41609 GCCCCGACGACGACTCACTACCCTCGGATGTGCCTATTGCTGTCGCACACCGCGCCOAC 41668

OY   208 TTCAGGGGGGAGCGGGCCCACTGTGTGGCAGCAGCCCGACAGGTGTGGGCGGACGCTGAT 267
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    41669 TTGAGCAGCCGGCGGCTCTGTGTAGCCCAACAGCGGACGAGCTCTCTCGCGCTCGAC 41728

OY   268 TTTATGCTCAAGGTCAA-AGAACCGATAGCGCGGAATAGCGCGCTCGACAGCGGGCG 326

```


FILED DATE: 23-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 18740/125/CACO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2185 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 531..2069
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 531..902
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 903..2069
FEATURE:
NAME/KEY: misc_feature
LOCATION: 531..533
OTHER INFORMATION: /note="Met at position -124
OTHER INFORMATION: represents fMet"
US-08-173-508-3

Query Match 3.4%; Score 42.4; DB 1; Length 2185;
Best Local Similarity 47.4%; Pred. No. 0.094;
Matches 127; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 524 CCGCGGTGCTGATGCGGGGGTCCCGGCGTCAACCGCGCGCATGCTGTGATCGG 583
DB 1250 CCGCGGTGCTGATGCGGGGGTCCCGGCGTCAACCGCGCGCATGCTGTGATCGG 1309
QY 584 CCGCGGACCGCGCGGCTTACACCGACCGCGCATGCGCGCAACGGCGCGCGCATGAC 643
DB 1310 CCGCGGACCGCGGCTTACACCGACCGCGCATGCGCGCAACGGCGCGCGCATGAC 1369
QY 644 GGTTCAGACATCAACATGACAACTTCGGCACTCGACGCGCGCATGCTGTGATCGG 703
DB 1370 CCGCGGTGCTGATGCGGGGGTCCCGGCGTCAACCGCGCGCATGCTGTGATCGG 1429
QY 704 CCGCGGTGCTGATGCGGGGGTCCCGGCGTCAACCGCGCGCATGCTGTGATCGG 763
DB 1430 CCGCGGTGCTGATGCGGGGGTCCCGGCGTCAACCGCGCGCATGCTGTGATCGG 1489
QY 764 GATTGGGGCGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 791
DB 1490 GAAGGGCGGGGTCAACGTCGCGCGCGCG 1517

RESULT 15
US-08-265-310-3
Sequence 3, Application US/08265310
Patent No. 5856166
GENERAL INFORMATION:
APPLICANT: Bartfeld, Daniel J.
APPLICANT: Butler, Michael J.
APPLICANT: Hadary, Dany
APPLICANT: Jenish, David
APPLICANT: Krieger, Timothy
APPLICANT: Malek, Lawrence T.
APPLICANT: Soostmeyer, Gisela
APPLICANT: Walczyk, Eva
APPLICANT: Kryzman, Phyllis
APPLICANT: Garven, Shella

TITLE OF INVENTION: STREPTOMYCES PROTEASES AND IMPROVED
TITLE OF INVENTION: STREPTOMYCES STRAINS FOR EXPRESSION OF PEPTIDES AND
TITLE OF INVENTION: POLYPEPTIDES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/265,310
FILING DATE: 24-JUN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,508
FILING DATE: 23-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 18740/133/CACO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2185 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 531..2069
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 531..902
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 903..2069
FEATURE:
NAME/KEY: misc_feature
LOCATION: 531..533
OTHER INFORMATION: /note="Met at position -124
OTHER INFORMATION: represents fMet"
US-08-265-310-3

Query Match 3.4%; Score 42.4; DB 3; Length 2185;
Best Local Similarity 47.4%; Pred. No. 0.094;
Matches 127; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 524 CCGCGGTGCTGATGCGGGGGTCCCGGCGTCAACCGCGCGCATGCTGTGATCGG 583
DB 1250 CCGCGGTGCTGATGCGGGGGTCCCGGCGTCAACCGCGCGCATGCTGTGATCGG 1309
QY 584 CCGCGGACCGCGCGGCTTACACCGACCGCGCATGCGCGCAACGGCGCGCGCATGAC 643
DB 1310 CCGCGGACCGCGGCTTACACCGACCGCGCATGCGCGCAACGGCGCGCGCATGAC 1369
QY 644 GGTTCAGACATCAACATGACAACTTCGGCACTCGACGCGCGCATGCTGTGATCGG 703
DB 1370 CCGCGGTGCTGATGCGGGGGTCCCGGCGTCAACCGCGCGCATGCTGTGATCGG 1429
QY 704 CCGCGGTGCTGATGCGGGGGTCCCGGCGTCAACCGCGCGCATGCTGTGATCGG 763
DB 1430 CCGCGGTGCTGATGCGGGGGTCCCGGCGTCAACCGCGCGCATGCTGTGATCGG 1489

Fri Jun 23 09:31:34 2000

us-09-362-485-6.rni

Page 12

Oy 764 GATTGGGCGCTCTGTGCCAGGC 791
||| ||| ||| |||
Db 1490 GAAGGCGCGTCAACGTGCGCGGCC 1517

Search completed: June 22, 2000, 15:11:10
Job time: 18148 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compygen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 12:07:36 ; Search time 5541.94 Seconds
(without alignments)
903.245 Million cell updates/sec

Title: US-09-362-485-6
Perfect score: 1235
Sequence: 1 ATCTTCAGATTAATCGAAC.....GGCAGCGATGATGCGGCC 1235

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 202611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: gb_est33:*
53: gb_est34:*
54: gb_est35:*
55: gb_est36:*
56: gb_est37:*
57: gb_est38:*
58: gb_est39:*
59: gb_est40:*
60: gb_est41:*
61: gb_est42:*
62: gb_est43:*
63: gb_est44:*
64: gb_est45:*
65: gb_est46:*
66: gb_est47:*
67: gb_est48:*
68: gb_est49:*
69: gb_est50:*
70: gb_est51:*
71: gb_est52:*
72: gb_est53:*
73: gb_est54:*
74: gb_est55:*
75: gb_est56:*
76: gb_est57:*
77: gb_est58:*
78: gb_est59:*
79: gb_est60:*
80: gb_est61:*
81: gb_est62:*
82: gb_est63:*
83: gb_est64:*
84: gb_est65:*
85: gb_est66:*
86: gb_est67:*
87: gb_est68:*
88: gb_est69:*
89: gb_est70:*
90: gb_est71:*
91: gb_est72:*
92: gb_est73:*
93: gb_est74:*
94: gb_est75:*
95: gb_est76:*
96: gb_est77:*
97: gb_est78:*
98: gb_est79:*
99: gb_est80:*
100: gb_est81:*
101: gb_est82:*
102: gb_est83:*
103: gb_est84:*
104: gb_est85:*
105: gb_est86:*
106: gb_est87:*
107: gb_est88:*
108: gb_est89:*
109: gb_est90:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result

Query

GenCore version 4.5
Copyright (c) 1993 - 2000 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 18:04:14 ; Search time 123.79 Seconds

(Without alignments)
15.751 Million cell updates/sec

Title: US-09-362-485-28

Sequence: 1 AACGATTCGGGTG 15

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued Patents, NA:*

1: /cgn2_6/ptodata/2/1na/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/1na/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/1na/5C_COMB.seq:*
4: /cgn2_6/ptodata/2/1na/5D_COMB.seq:*
5: /cgn2_6/ptodata/2/1na/6_COMB.seq:*
6: /cgn2_6/ptodata/2/1na/PCITUS_COMB.seq:*
7: /cgn2_6/ptodata/2/1na/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	13.4	89.3	206	US-08-644-034A-5	Sequence 5, Appli
2	13	86.7	80	US-08-468-275-2	Sequence 2, Appli
3	13	86.7	207	US-08-142-551B-8	Sequence 8, Appli
4	12.4	82.7	48	US-09-211-631-20	Sequence 20, Appli
5	12.4	82.7	153	US-08-651-136C-49	Sequence 49, Appli
6	12.4	82.7	1389	US-08-702-344-20	Sequence 20, Appli
7	12.4	82.7	4145	US-08-314-917-1	Sequence 1, Appli
8	12.4	82.7	4145	US-08-265-046-1	Sequence 1, Appli
9	12.4	82.7	4145	US-08-465-522-1	Sequence 1, Appli
10	12.4	82.7	4145	US-08-140-11401-1	Sequence 1, Appli
11	12.4	82.7	4145	US-08-173-510B-1	Sequence 1, Appli
12	12.4	82.7	35	US-08-458-218-51	Sequence 51, Appli
13	12.4	82.7	35	US-08-459-457-51	Sequence 51, Appli
14	12.4	82.7	35	US-08-540-406-1	Sequence 1, Appli
15	12.4	82.7	736	US-08-656-055-1	Sequence 1, Appli
16	12.4	82.7	736	US-08-656-055-1	Sequence 1, Appli
17	12.4	82.7	736	US-08-656-055-1	Sequence 1, Appli
18	12.4	82.7	1699	US-08-484-933B-3	Sequence 3, Appli
19	12.4	82.7	1699	US-08-484-933B-3	Sequence 3, Appli
20	12.4	82.7	1699	US-08-484-933B-3	Sequence 3, Appli
21	12.4	82.7	1699	US-08-484-933B-3	Sequence 3, Appli
22	12.4	82.7	1699	US-08-484-933B-3	Sequence 3, Appli
23	12.4	82.7	1699	US-08-484-933B-3	Sequence 3, Appli
24	12.4	82.7	1699	US-08-484-933B-3	Sequence 3, Appli
25	11.8	78.7	171	US-08-894-440-4	Sequence 4, Appli
26	11.8	78.7	224	US-08-644-034A-4	Sequence 4, Appli
27	11.8	78.7	326	US-08-111-316-3	Sequence 3, Appli

28	11.8	78.7	326	1	US-08-111-316-5	Sequence 5, Appli
29	11.8	78.7	326	1	US-08-468-405-3	Sequence 3, Appli
30	11.8	78.7	326	1	US-08-468-405-3	Sequence 3, Appli
31	11.8	78.7	348	3	US-08-642-229A-10	Sequence 10, Appli
32	11.8	78.7	375	1	US-08-111-316-4	Sequence 4, Appli
33	11.8	78.7	375	1	US-08-468-405-4	Sequence 4, Appli
34	11.8	78.7	429	5	US-08-520-550A-47	Sequence 47, Appli
35	11.8	78.7	476	1	US-08-111-316-2	Sequence 2, Appli
36	11.8	78.7	476	1	US-08-468-405-2	Sequence 2, Appli
37	11.8	78.7	560	5	US-08-520-550A-46	Sequence 46, Appli
38	11.8	78.7	580	3	US-08-485-778-45	Sequence 45, Appli
39	11.8	78.7	580	3	US-08-520-550A-43	Sequence 43, Appli
40	11.8	78.7	605	3	US-08-769-967A-33	Sequence 33, Appli
41	11.8	78.7	1128	2	US-08-815-688A-4	Sequence 4, Appli
42	11.8	78.7	1246	1	US-08-446-777-3	Sequence 3, Appli
43	11.8	78.7	1312	1	US-08-205-506A-1	Sequence 1, Appli
44	11.8	78.7	1312	6	PCT-US94-02389-1	Sequence 1, Appli
45	11.8	78.7	1313	3	US-08-463-911-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-08-644-034A-5
Sequence 5, Application US/08644034A
Patent No. 5882862
GENERAL INFORMATION:
APPLICANT: Howard B. Lieberman and Kevin M. Hopkins
TITLE OF INVENTION: A HUMAN RADIORRESISTANCE/CELL CYCLE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham, LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/644,034A
FILING DATE: May 9, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: John P. White
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/50307
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0526
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 206 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-08-644-034A-5

Query Match 89.3% Score 13.4; DB 3; Length 206;
Best Local Similarity 93.3%; Pred. NO. 29;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AACGATTCGGGTG 15
DB 44 AACGATTCGGGTG 58

```
RESULT 2
US-08-468-275-2
; Sequence 2, Application US/08468275
; Patent No. 5747453
; GENERAL INFORMATION:
; APPLICANT: HOLLDAY, LESLIE A.
; APPLICANT: OLDENBURG, KEVIN R.
; TITLE OF INVENTION: METHOD FOR INCREASING THE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ALZA CORPORATION
; STREET: 950 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94303-0802
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,275
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: MILLER, D. BYRON
; REGISTRATION NUMBER: 30,661
; REFERENCE/DOCKET NUMBER: 0360-0002; ARC-2349
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 496-8150
; TELEFAX: (415) 496-8048
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-468-275-2

Query Match      86.7%; Score 13; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ACGAATCCGGGT 14
    |||||||
Db 59 ACGAATCCGGGT 71

RESULT 3
US-08-142-551B-8/c
; Sequence 8, Application US/08142551B
; Patent No. 5814603
; GENERAL INFORMATION:
; APPLICANT: Oldenburg, Kevin R.
; APPLICANT: Selick, Harold E.
; TITLE OF INVENTION: COMPOUNDS WITH PTH ACTIVITY AND
; NUMBER OF SEQUENCES: 132
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: US
; ZIP: 22313
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/142,551B
FILING DATE: 25-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/077,296
FILING DATE: 14-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/898,219
FILING DATE: 12-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/965,677
FILING DATE: 22-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Swiss, Gerald F.
REGISTRATION NUMBER: 30,113
REFERENCE/DOCKET NUMBER: 000324-010
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 854-7400
TELEFAX: (415) 854-8275
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 207 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..201
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..51
OTHER INFORMATION: /note="Encodes the leader peptide
OTHER INFORMATION: sequence that serves to direct the protein into
OTHER INFORMATION: inclusion bodies."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 70..174
OTHER INFORMATION: /note="Encodes the protein or
OTHER INFORMATION: peptide of interest"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 175..207
OTHER INFORMATION: /note="Encodes amino acid sequence
OTHER INFORMATION: having six histamines that serves as a tag for the
OTHER INFORMATION: purification of the protein on a nickel column."
US-08-142-551B-8

Query Match      86.7%; Score 13; DB 2; Length 207;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ACGAATCCGGGT 14
    |||||||
Db 53 ACGAATCCGGGT 41

RESULT 4
US-09-211-631-20/c
; Sequence 20, Application US/09211631
; Patent No. 6001597
; GENERAL INFORMATION:
; APPLICANT: Raymond, Christopher K.
; APPLICANT: Vanaia, Erica
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PRODUCING
; HETEROLOGOUS POLYPEPTIDES IN PICHIA METHANOLICA
; FILE REFERENCE: 96-16C2
; CURRENT APPLICATION NUMBER: US/09/211,631
; CURRENT FILING DATE: 1998-12-15
```

NUMBER OF SEQ ID NOS: 23
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 20
LENGTH: 48
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: oligonucleotide primer
US-09-211-631-20

Query Match 82.7%; Score 12.4; DB 5; Length 48;
Best Local Similarity 92.9%; Pred. No. 1e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACGAATTCGGGT 14
DB 21 ACGAATTCGGAT 8

RESULT 5

US-08-651-136C-49
Sequence 49, Application US/08651136C
Patent No. 6001639

GENERAL INFORMATION:
APPLICANT: Schuelein, Martin
APPLICANT: Andersen, Lene N.
APPLICANT: Lassen, Soren F.
APPLICANT: Kauppinen, Markus S.
APPLICANT: Lange, Lene
APPLICANT: Nielsen, Ruby I.
APPLICANT: Ihara, Michiko
APPLICANT: Takagi, Shinobu
TITLE OF INVENTION: No. 6001639e1 Endoglucanases
NUMBER OF SEQUENCES: 109
CORRESPONDENCE ADDRESSES:
ADDRESSEE: No. 6001639c No. 6001639dsk of No. 6001639th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/651,136C
FILING DATE: 21-MAY-1996
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728

REFERENCE/DOCKET NUMBER: 4366.200-US
TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 49:

SEQUENCE CHARACTERISTICS:
LENGTH: 153 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: cDNA

FEATURE:
NAME/KEY: CDS
LOCATION: 1..153

US-08-651-136C-49

Query Match 82.7%; Score 12.4; DB 5; Length 153;
Best Local Similarity 92.9%; Pred. No. 1.1e+02;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ACGAATTCGGGTG 15
DB 37 ACGAATTCGGGTG 50

RESULT 6

US-08-702-344-20
Sequence 20, Application US/08702344
Patent No. 5723315

GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John
APPLICANT: Lavallee, Edward
APPLICANT: Racie, Lisa
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,344
FILING DATE:
CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:
LENGTH: 1389 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

TOPOLOGY: linear
MOLECULE TYPE: cDNA

US-08-702-344-20

Query Match 82.7%; Score 12.4; DB 2; Length 1389;
Best Local Similarity 92.9%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ACGAATTCGGGTG 15
DB 431 ACGAATTCGGGTG 444

RESULT 7

US-08-314-917-1
Sequence 1, Application US/08314917
Patent No. 5468630

GENERAL INFORMATION:
APPLICANT: Billiar, Timothy R.
APPLICANT: Nussler, Andreas K.
APPLICANT: Geller, David A.
APPLICANT: Simmons, Richard L.
TITLE OF INVENTION: cDNA Clone for Human Inducible Nitric

TITLE OF INVENTION: Oxide Synthase And Process for Preparing Same
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold B. Silverman
ADDRESSEE: Eckert Seamans Cherin & Mellott
STREET: 600 Grant Street, 42nd Floor
CITY: Pittsburgh
STATE: PA
COUNTRY: USA
ZIP: 15219
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/314,917
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/981,344
FILING DATE: 25-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Silverman, Arnold B.
REGISTRATION NUMBER: 22,614
REFERENCE/DOCKET NUMBER: 116972
TELECOMMUNICATION INFORMATION:
TELEPHONE: (412) 566-6000
TELEFAX: (412) 566-6099
TELEX: 866172
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4145 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
DESCRIPTION: Human Hepatocyte Inducible Nitric Oxide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
TISSUE TYPE: Induced Human Hepatocyte RNA
IMMEDIATE SOURCE:
LIBRARY: Lambda Zap II cDNA
CLONE: PHINOS
POSITION IN GENOME:
CHROMOSOME/SEGMENT: unknown
MAP POSITION: unknown
UNITS: unknown
FEATURE:
NAME/KEY: CDS
LOCATION: 207..3668
IDENTIFICATION METHOD: Experiment
US-08-314-917-1

Query Match 82.7%; Score 12.4; DB 1; Length 4145;
Best Local Similarity 92.9%; Pred. No. 1.1e+2;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ACGAATTCGGGGTG 15
|||||

DB 970 ACGACTTCGGGGTG 983

RESULT 8
US-08-265-046-1
; Sequence 1, Application US/08265046
; Patent No. 5658565
; GENERAL INFORMATION:
; APPLICANT: Timothy R. Billiar
; APPLICANT: Edith Tzeng

APPLICANT: Andreas K. Nussler
APPLICANT: David A. Geller
APPLICANT: Richard L. Simmons
TITLE OF INVENTION: Inducible Nitric Oxide Synthase
TITLE OF INVENTION: Gene for Treatment of Disease
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lewis F. Gould, Jr.
ADDRESSEE: Eckert Seamans Cherin & Mellott
STREET: 1700 Market Street, Suite 3232
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/265,046
FILING DATE: 24-JUN-1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Gould, Lewis F. Jr.
REGISTRATION NUMBER: 25,057
REFERENCE/DOCKET NUMBER: 119130
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 575-6020
TELEFAX: (215) 575-6015
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4145 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
DESCRIPTION: Human Hepatocyte Inducible Nitric Oxide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
TISSUE TYPE: Induced Human Hepatocyte RNA
IMMEDIATE SOURCE:
LIBRARY: Lambda Zap II cDNA
CLONE: PHINOS
POSITION IN GENOME:
CHROMOSOME/SEGMENT: unknown
MAP POSITION: unknown
UNITS: unknown
FEATURE:
NAME/KEY: CDS
LOCATION: 207..3668
IDENTIFICATION METHOD: Experiment
US-08-265-046-1

Query Match 82.7%; Score 12.4; DB 1; Length 4145;
Best Local Similarity 92.9%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ACGAATTCGGGGTG 15
|||||

DB 970 ACGACTTCGGGGTG 983

RESULT 9
US-08-465-522-1
; Sequence 1, Application US/08465522
; Patent No. 5882908
; GENERAL INFORMATION:
; APPLICANT: Billiar, Timothy R.

APPLICANT: Nusaler, Andreas K.
APPLICANT: Geller, David A.
APPLICANT: Simmons, Richard L.
TITLE OF INVENTION: CDNA Clone for Human Inducible Nitric
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lewis F. Gould, Jr.
ADDRESS: Eckert Seamans Cherin & Mellott
STREET: 1700 Market St. Suite 3232
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,522
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gould, Jr., Lewis F.
REGISTRATION NUMBER: 25,057
REFERENCE/DOCKET NUMBER: 116972-6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 575-6020
TELEFAX: (215) 575-6015
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4145 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
DESCRIPTION: Human Hepatocyte Inducible Nitric Oxide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
TISSUE TYPE: Induced Human Hepatocyte RNA
IMMEDIATE SOURCE:
LIBRARY: Lambda Zap II CDNA
CLONE: PHINOS
POSITION IN GENOME:
CHROMOSOME/SEGMENT: unknown
MAP POSITION: unknown
UNITS: unknown
FEATURE:
NAME/KEY: CDS
LOCATION: 207..3668
IDENTIFICATION METHOD: Experiment
US-08-465-522-1

Query Match 82.7%; Score 12.4; DB 3; Length 4145;
Best Local Similarity 92.9%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ACGAATCCGGGTG 15
DB 970 ACGACTTCGGGTG 983

RESULT 10
PCT-US93-11401-1
Sequence 1, Application PCTUS9311401
GENERAL INFORMATION:
APPLICANT: Billiar, Timothy R.
APPLICANT: Nusaler, Andreas K.

APPLICANT: Geller, David A.
APPLICANT: Simmons, Richard L.
TITLE OF INVENTION: CDNA Clone for Human Inducible Nitric
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold B. Silverman
ADDRESS: Eckert Seamans Cherin & Mellott
STREET: 600 Grant Street, 42nd Floor
CITY: Pittsburgh
STATE: PA
COUNTRY: USA
ZIP: 15219
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/11401
FILING DATE: 25-NOV-1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/981,344
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Silverman, Arnold B.
REGISTRATION NUMBER: 22,614
REFERENCE/DOCKET NUMBER: 116972
TELECOMMUNICATION INFORMATION:
TELEPHONE: (412) 566-6000
TELEFAX: (412) 566-6099
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4145 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
TISSUE TYPE: Induced Human Hepatocyte RNA
IMMEDIATE SOURCE:
LIBRARY: Lambda Zap II CDNA
CLONE: PHINOS
POSITION IN GENOME:
CHROMOSOME/SEGMENT: unknown
MAP POSITION: unknown
UNITS: unknown
FEATURE:
NAME/KEY: CDS
LOCATION: 207..3668
IDENTIFICATION METHOD: Experiment
PCT-US93-11401-1

Query Match 82.7%; Score 12.4; DB 6; Length 4145;
Best Local Similarity 92.9%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ACGAATCCGGGTG 15
DB 970 ACGACTTCGGGTG 983

RESULT 11
PCT-US95-07849-1
Sequence 1, Application PC/TUS9507849
GENERAL INFORMATION:
APPLICANT: University of Pittsburgh of the Commonwealth System of Higher
APPLICANT: Education

TITLE OF INVENTION: Inducible Nitric Oxide Synthase
TITLE OF INVENTION: Gene for Treatment of Disease
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lewis F. Gould, Jr.
ADDRESSEE: Eckert Seamans Cherin & Mellott
STREET: 1700 Market Street, Suite 3232
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07849
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Gould, Lewis F. Jr.
REGISTRATION NUMBER: 25,057
REFERENCE/DOCKET NUMBER: 119130-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 575-6020
TELEFAX: (215) 575-6015
TELEX:
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4145 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
DESCRIPTION: Human Hepatocyte Inducible Nitric Oxide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
TISSUE TYPE: Induced Human Hepatocyte RNA
IMMEDIATE SOURCE:
LIBRARY: Lambda Zap II cDNA
CLONE: PHINOS
POSITION IN GENOME:
CHROMOSOME/SEGMENT: unknown
MAP POSITION: unknown
UNITS: unknown
FEATURE:
NAME/KEY: CDS
LOCATION: 207..3668
IDENTIFICATION METHOD: Experiment
PCT-US95-07849-1

Query Match 82.7%; Score 12.4; DB 6; Length 4145;
Best Local Similarity 92.9%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ACGAATCCGGCGTG 15
DB 970 ACGAATCCGGCGTG 983

RESULT 12
US-08-173-510B-51/c
Sequence 51, Application US/08173510B
Patent No. 5747296
GENERAL INFORMATION:
APPLICANT: MATTHEW MOYLE ET AL.
TITLE OF INVENTION: NOVEL NEUTROPHIL INHIBITORS
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: Storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/173,510B
FILING DATE: 23-DEC-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/151,064
FILING DATE: 10-NOV-1993
APPLICATION NUMBER: 08/060,433
FILING DATE: 11-MAY-1993
APPLICATION NUMBER: 07/996,972
FILING DATE: 24-DEC-1992
APPLICATION NUMBER: 07/881,721
FILING DATE: 11-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 205/073
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ. ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 35
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
MOLECULE TYPE: NUCLEIC

US-08-173-510B-51

Query Match 80.0%; Score 12; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGAATCCGG 12
DB 12 ACGAATCCGG 1

RESULT 13
US-08-458-218-51/c
Sequence 51, Application US/08458218
Patent No. 5789178
GENERAL INFORMATION:
APPLICANT: MATTHEW MOYLE ET AL.
TITLE OF INVENTION: NOVEL NEUTROPHIL INHIBITORS
NUMBER OF SEQUENCES: 102
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: Storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,218
FILING DATE: 11-MAY-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/151,064
FILING DATE: 10-NOVEMBER-1993
APPLICATION NUMBER: 08/060,433
FILING DATE: 11-MAY-1993
APPLICATION NUMBER: 07/881,721
FILING DATE: 11-MAY-1992
APPLICATION NUMBER: 07/996,972
FILING DATE: 24-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 203/726
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 35
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
MOLECULE TYPE: NUCLEIC
US-08-458-218-51

Query Match 80.0%; Score 12; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACGAATCCCG 12
|||||
DB 12 AACGAATCCCG 1

RESULT 14
US-08-450-497-51/C
Sequence 51, Application US/08450497
Patent No. 5919900
GENERAL INFORMATION:
APPLICANT: MATTHEW MOYLE, ET AL.
TITLE OF INVENTION: NOVEL NEUTROPHIL INHIBITORS
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
STATE: Los Angeles
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,497
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/173,510
FILING DATE: 23-DEC-1993
APPLICATION NUMBER: 08/151,064
FILING DATE: 10-NOV-1993
APPLICATION NUMBER: 08/060,433
FILING DATE: 11-MAY-1993

APPLICATION NUMBER: 07/996,972
FILING DATE: 24-DEC-1992
APPLICATION NUMBER: 07/881,721
FILING DATE: 11-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 205/073
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 35
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
MOLECULE TYPE: NUCLEIC
US-08-450-497-51

Query Match 80.0%; Score 12; DB 4; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACGAATCCCG 12
|||||
DB 12 AACGAATCCCG 1

RESULT 15
US-08-540-406-1/C
Sequence 1, Application US/08540406
Patent No. 5337538
GENERAL INFORMATION:
APPLICANT: SCOTT, MATHEW P
APPLICANT: GOODRICH, LISA V
APPLICANT: JOHNSON, RONALD L
TITLE OF INVENTION: Patched Genes and their use
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/540,406
FILING DATE: 06-OCT-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: 20015
REFERENCE/DOCKET NUMBER: a60190-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 736 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-540-406-1

Query Match 80.0%; Score 12; DB 3; Length 736;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0;

QY 4 GAATTCGGGTG 15
 |||||

Db 147 GAATTCGGGTG 136

Search completed: June 22, 2000, 18:04:15
 Job time: 9722 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 14:50:07 ; Search time 8627.09 Seconds

(Without alignments)
-140.386 Million cell updates/sec

Title: US-09-362-485-2
Perfect score: 1245
Sequence: 1 ATCTGCAGATTAATCGAAC.....AAGCGATGATGCGCGCG 1245

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 segs, -486395729 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenBankl.*
1: gb_ba1.*
2: gb_ba2.*
3: gb_om.*
4: gb_om.*
5: gb_om.*
6: gb_ph.*
7: gb_pl1.*
8: gb_pl2.*
9: gb_pl3.*
10: gb_pl3.*
11: gb_pl3.*
12: gb_pl3.*
13: gb_pl3.*
14: gb_pl3.*
15: gb_pl3.*
16: gb_pl3.*
17: gb_pl3.*
18: gb_pl3.*
19: gb_pl3.*
20: gb_pl3.*
21: gb_pl3.*
22: gb_pl3.*
23: gb_pl3.*
24: gb_pl3.*
25: gb_pl3.*
26: gb_pl3.*
27: gb_pl3.*
28: gb_pl3.*
29: gb_pl3.*
30: gb_pl3.*
31: gb_pl3.*
32: gb_pl3.*
33: gb_pl3.*
34: gb_pl3.*
35: gb_pl3.*
36: gb_pl3.*
37: gb_pl3.*
38: gb_pl3.*
39: gb_pl3.*
40: gb_pl3.*
41: gb_pl3.*
42: gb_pl3.*
43: gb_pl3.*
44: gb_pl3.*

45: gb_hlg7.*
46: em_hlg1.*
47: em_hlg2.*
48: em_hlg3.*
49: em_hlg4.*
50: gb_pl3.*
51: gb_pl3.*
52: gb_pl3.*
53: gb_pl3.*
54: gb_pl3.*
55: gb_pl3.*
56: gb_pl3.*
57: gb_pl3.*
58: gb_hlg14.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1245	100.0	1245	5 A87605	A87605 Sequence 2
2	1245	100.0	1245	5 A89745	A89745 Sequence 2
3	1245	100.0	1260	5 A87604	A87604 Sequence 1
4	1245	100.0	1260	5 A89744	A89744 Sequence 1
5	1245	100.0	2584	1 MTVALDH	X63069 M. tuberculo
6	1221.4	98.1	56414	1 MTVALDH	AL008967 Mycobacte
7	1220	98.0	1235	5 A87613	A87613 Sequence 10
8	1220	98.0	1235	5 A89753	A89753 Sequence 10
9	1219.4	97.9	1237	5 A87607	A87607 Sequence 4
10	1219.4	97.9	1237	5 A89747	A89747 Sequence 4
11	1216	97.7	1235	5 A87606	A87606 Sequence 3
12	1216	97.7	1235	5 A89746	A89746 Sequence 3
13	1210.4	97.2	1228	5 A87608	A87608 Sequence 5
14	1210.4	97.2	1228	5 A89748	A89748 Sequence 5
15	1207	96.9	1235	5 A87609	A87609 Sequence 6
16	1207	96.9	1235	5 A89749	A89749 Sequence 6
17	1206.4	96.9	1235	5 A87611	A87611 Sequence 8
18	1206.4	96.9	1235	5 A89751	A89751 Sequence 8
19	1199.4	96.3	1229	5 A87610	A87610 Sequence 7
20	1199.4	96.3	1229	5 A89750	A89750 Sequence 7
21	1192	95.7	1209	5 A87612	A87612 Sequence 9
22	1185	95.2	1208	5 A89752	A89752 Sequence 9
23	1178	94.6	1194	2 MTU92472	U92472 Mycobacteri
24	720	57.8	720	5 A89759	A89759 Sequence 16
25	522.4	42.0	40745	1 SC151	AL109848 Streptomy
26	467.6	37.6	682	5 A89758	A89758 Sequence 15
27	467.6	37.6	682	5 A89762	A89762 Sequence 19
28	361	29.0	8046	2 AF049107	AF049107 Myxococcu
29	337.2	27.1	11514	2 AE002028	AE002028 Deinococc
30	315.4	25.3	2780	1 RLE238118	AJ238118 Rhizobium
31	314.8	25.3	1585	2 AF070716	AF070716 Vibrio pr
32	296.8	23.8	1736	1 BACALDH	M33299 B. steatoche
33	277	22.2	1477	2 AF070715	AF070715 Shewanell
34	272.4	21.9	1123	1 D37807	D37807 Phormidium
35	261.2	21.0	1125	5 E01847	E01847 DNA encodin
36	246.8	19.8	13740	1 D90900	D90900 Synecocyst
37	227.2	18.2	1375	1 AB013821	AB013821 Enterobac
38	227	18.2	1776	1 BACALDH	M33298 B. sphaericu
39	219.6	17.6	1709	1 BACALDH	L20916 Bacillus su
40	219.6	17.6	12316	1 BS282015	BS282015 B. subtilis
41	219.6	17.6	217420	1 BSUB0017	Z99120 Bacillus su
42	202.2	16.2	1714	2 AF070714	AF070714 Carnobact
43	108.8	8.7	10652	2 AE001565	AE001565 Helicobac
44	100.8	8.1	10512	2 AE000639	AE000639 Helicobac
45	66.4	5.3	3160	1 RR005294	U05294 Rhodospirill

ALIGNMENTS

```

RESULT 1
LOCUS A87605 1245 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 2 from Patent WO9836089.
ACCESSION A87605
VERSION A87605.1 GI:6736245
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1245)
AUTHORS Flohe, L. and Singh, M.
TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL Patent: WO 9836089-A-20-AUG-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
Source 1..1245
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 238 a 398 c 387 g 222 t
ORIGIN
Query Match 100.0%; Score 1245; DB 5; Length 1245;
Best Local Similarity 100.0%; Pred. No. 1.2e-171;
Matches 1245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTTGAGATTATCGAACTTCTTCACACTGAGAGGTACAGTATGAGAGTGTATATC 60
DB 1 ATCTTGAGATTATCGAACTTCTTCACACTGAGAGGTACAGTATGAGAGTGTATATC 60
QY 61 ATGGCGGTGCTATTCGACCGAGCAACAAAAACAAGAAATTCATTCGGGTGGCCATC 120
DB 61 ATGGCGGTGCTATTCGACCGAGCAACAAAAACAAGAAATTCATTCGGGTGGCCATC 120
QY 121 ACCCGCGCGCGCTGCGGGAAGTAAACCGGTGTCGATGAGTGTCTATCCAGGAGGT 180
DB 121 ACCCGCGCGCGCTGCGGGAAGTAAACCGGTGTCGATGAGTGTCTATCCAGGAGGT 180
QY 181 GCGGAGAGGGCTGGGCTATCAGCGAGCGGATTCAGAGGGGAGGGGCGCACTGGTC 240
DB 181 GCGGAGAGGGCTGGGCTATCAGCGAGCGGATTCAGAGGGGAGGGGCGCACTGGTC 240
QY 241 GGCACCGCGCGAGCAAGTGTGGCGGAGCGCTGATTTATGCTCAAGGTCAAGAAACGATA 300
DB 241 GGCACCGCGCGAGCAAGTGTGGCGGAGCGCTGATTTATGCTCAAGGTCAAGAAACGATA 300
QY 301 GCGGCGGAATACGGCGCGCTGCGACAGCGGAGATCTTGTTCACAGTCTTGGCATTTGGCC 360
DB 301 GCGGCGGAATACGGCGCGCTGCGACAGCGGAGATCTTGTTCACAGTCTTGGCATTTGGCC 360
QY 361 GCGTACAGTGTTCACAGTATGCTGTTGATTCGCGACACCGTCAATGGCTTAGAG 420
DB 361 GCGTACAGTGTTCACAGTATGCTGTTGATTCGCGACACCGTCAATGGCTTAGAG 420
QY 421 ACCGTCCAGACCGCGGAGCGGAGTACCCCTGCTTGGCCCGATGAGGAAATGCGCGGT 480
DB 421 ACCGTCCAGACCGCGGAGCGGAGTACCCCTGCTTGGCCCGATGAGGAAATGCGCGGT 480
QY 481 CGAATCGCGCGCGAGTGTGGGCTTACCACTGATGGAAACCAAGGGGCGCGGTGTG 540
DB 481 CGAATCGCGCGCGAGTGTGGGCTTACCACTGATGGAAACCAAGGGGCGCGGTGTG 540
QY 541 CTGATGGGCGGTGTGCCGCGGTGGAACCGGCGACGTCGTGTGATCGGCGCGGAC 600
DB 541 CTGATGGGCGGTGTGCCGCGGTGGAACCGGCGACGTCGTGTGATCGGCGCGGAC 600
QY 601 GCGGCGTACAAAGCGAGCGCGCATGCCAAAGCGCATGGGCGGAGCGTTACGTTAGAC 660
DB 601 GCGGCGTACAAAGCGAGCGCGCATGCCAAAGCGCATGGGCGGAGCGTTACGTTAGAC 660
QY 661 ATCAACATCGAACAATTCGGCACTGAGCGCCAGATTTTCGGCGCGGATCCACACTGCC 720

```

```

DB 661 ATCAACATCGAACAATTCGGCACTGAGCGCCGAGTTCGCGCGCGGATCCACACTGCC 720
QY 721 TACTCATGCGCTTACGAGCTGAGGGTCCCGTAAAGTCCGACCTGTGATTTGGGGCC 780
DB 721 TACTCATGCGCTTACGAGCTGAGGGTCCCGTAAAGTCCGACCTGTGATTTGGGGCC 780
QY 781 GTCTGTGCGAGCGCGCAAGGCAACCAATTAGTCTGAAATTCATTTGCGGCAATAG 840
DB 781 GTCTGTGCGAGCGCGCAAGGCAACCAATTAGTCTGAAATTCATTTGCGGCAATAG 840
QY 841 AAACAGTCCGCTACGCTGATATAGCATGACAGGCGGCTGTTTCGAAGGCTCA 900
DB 841 AAACAGTCCGCTACGCTGATATAGCATGACAGGCGGCTGTTTCGAAGGCTCA 900
QY 901 CGACCGACCACTTACGACCAACCCGAGCTCCGCTGACAGACAGCTGTTTACTGCTG 960
DB 901 CGACCGACCACTTACGACCAACCCGAGCTTCGCGTACAGACAGCTGTTTACTGCTG 960
QY 961 GCGAATGCGCGCGCTGCGGAGAGTGCACCTAGCGGCTGACCAACGAGAGATG 1020
DB 961 GCGAATGCGCGCGCTGCGGAGAGTGCACCTAGCGGCTGACCAACGAGAGATG 1020
QY 1021 CCGTATGCTCGAGCTTCCGACCATGCTGCGGCGCGGCGGTGCGATTCGATCCGCA 1080
DB 1021 CCGTATGCTCGAGCTTCCGACCATGCTGCGGCGCGGCGGTGCGATTCGATCCGCA 1080
QY 1081 CTAGCCAAAGTCTTTCGAGCGACGAAAGGGCGTTACTGTCACAAAGGGTGGCCACCGAC 1140
DB 1081 CTAGCCAAAGTCTTTCGAGCGACGAAAGGGCGTTACTGTCACAAAGGGTGGCCACCGAC 1140
QY 1141 CTGGGGGTGCGCTTTCACCGAGCGCGGAGGCTGCTGCGCTGATCTTGGCGGCTGTAC 1200
DB 1141 CTGGGGGTGCGCTTTCACCGAGCGCGGAGGCTGCTGCGCTGATCTTGGCGGCTGTAC 1200
QY 1201 GCGGAGCACAGTGGGAGTAAGGAAAGCATGATGTGCGCCGCG 1245
DB 1201 GCGGAGCACAGTGGGAGTAAGGAAAGCATGATGTGCGCCGCG 1245

RESULT 2
LOCUS A89745 1245 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 2 from Patent WO9832862.
ACCESSION A89745
VERSION A89745.1 GI:6738279
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1245)
AUTHORS Flohe, L. and Singh, M.
TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
JOURNAL Patent: WO 9832862-A-30-JUL-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
Source 1..1245
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 238 a 398 c 387 g 222 t
ORIGIN
Query Match 100.0%; Score 1245; DB 5; Length 1245;
Best Local Similarity 100.0%; Pred. No. 1.2e-171;
Matches 1245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTTGAGATTATCGAACTTCTTCACACTGAGAGGTACAGTATGAGAGGTTATC 60
DB 1 ATCTTGAGATTATCGAACTTCTTCACACTGAGAGGTACAGTATGAGAGGTTATC 60
QY 61 ATGGCGGTGCTATTCGACCGAGCAACAAAAACAAGAAATTCATTCGGGTGGCCATC 120

```

```

Db 61 ATGGCGTGGTATTCGACCGAGACCAAAACAAGATTCAATTCCGGGTGGCCATC 120
QY 121 ACCCGCGCGCGCTGGGGAATTAACCCGTCGTGGCCATGAGTGTCTCATTCAGGAGGT 180
Db 121 ACCCGCGCGCGCTGGGGAATTAACCCGTCGTGGCCATGAGTGTCTCATTCAGGAGGT 180
QY 181 GCGGAGAGGGCTGGGCTATACCGAGCGGATTTCAAGGGGGGCGGCGCACTGTGC 240
Db 181 GCGGAGAGGGCTGGGCTATACCGAGCGGATTTCAAGGGGGGCGGCGCACTGTGC 240
QY 241 GGCACCGCGCGAGAGTGTGGGCGGAGCTGATTTATTTCTCAAGGTCAAGAACGATA 300
Db 241 GGCACCGCGCGAGAGTGTGGGCGGAGCTGATTTATTTCTCAAGGTCAAGAACGATA 300
QY 301 GCGGCGGAATACGGCGCGCTGCGACAGGGGAGATCTTGTCTGACGTTCTTGCATTGGCC 360
Db 301 GCGGCGGAATACGGCGCGCTGCGACAGGGGAGATCTTGTCTGACGTTCTTGCATTGGCC 360
QY 361 GCGTCACGTGCTTGCACCGATGCTGTGTGATTCGGGACACGATGATTCCTAGAG 420
Db 361 GCGTCACGTGCTTGCACCGATGCTGTGTGATTCGGGACACGATGATTCCTAGAG 420
QY 421 ACCGTCAGACCGCGGAGCGGCGACATCCCTGCTTCCCGGATGAGCGAAGTCCGGGT 480
Db 421 ACCGTCAGACCGCGGAGCGGCGACATCCCTGCTTCCCGGATGAGCGAAGTCCGGGT 480
QY 481 CGAGTCGCGCGCGGAGTGGGGCTTACCACTGATGCGAACCCAAAGGGGGCGCGGTGTG 540
Db 481 CGAGTCGCGCGCGGAGTGGGGCTTACCACTGATGCGAACCCAAAGGGGGCGCGGTGTG 540
QY 541 CTGATGGGCGGGGTGCCCGCGGTGCGAACCGGCGCGGCGGTGTGATGCGCGCGCAC 600
Db 541 CTGATGGGCGGGGTGCCCGCGGTGCGAACCGGCGCGGCGGTGTGATGCGCGCGCAC 600
QY 601 GCGGCTTCAACGACGCGCGGCTATGCGCAAGCGGATGGGCGGAGCGGATTCAGTAC 660
Db 601 GCGGCTTCAACGACGCGCGGCTATGCGCAAGCGGATGGGCGGAGCGGATTCAGTAC 660
QY 661 ATCAACATCGAACAATCTGCGCACTGACGCGGATCTCGGCGCGGATTCACACTGCG 720
Db 661 ATCAACATCGAACAATCTGCGCACTGACGCGGATCTCGGCGCGGATTCACACTGCG 720
QY 721 TACTCATCGGCTTACGAGCTGAGGGTCCGTCAAAGTCCGACCTGTGTGATTTGGGCC 780
Db 721 TACTCATCGGCTTACGAGCTGAGGGTCCGTCAAAGTCCGACCTGTGTGATTTGGGCC 780
QY 781 GTCTGTGTGCGAGGGCGGCAAGGCAACCAATTAGTCTGGAATTCATTGTGCGCATATG 840
Db 781 GTCTGTGTGCGAGGGCGGCAAGGCAACCAATTAGTCTGGAATTCATTGTGCGCATATG 840
QY 841 AAACAGGTGGGCTACTGTGATATAGCATTCAGACAGGGCGGCTGTTTGAAGGCTCA 900
Db 841 AAACAGGTGGGCTACTGTGATATAGCATTCAGACAGGGCGGCTGTTTGAAGGCTCA 900
QY 901 CGACGACACCTTACGACACCAACCGAGTTCGCGCTGACACGACAGCTTTTACTGCTG 960
Db 901 CGACGACACCTTACGACACCAACCGAGTTCGCGCTGACACGACAGCTTTTACTGCTG 960
QY 961 GCGACATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Db 961 GCGACATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 1021 GCGGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db 1021 GCGATGTGCTGAGCTTGCAGACCATGCTGCGGGGGGCTGCGGCTGCGAATCCGGA 1080
QY 1081 CTAGCCAAAGTCTTTCAGACGACGAGAGGGGCTTACTGTCGAAAGGCTGCGCAC 1140
Db 1081 CTAGCCAAAGTCTTTCAGACGACGAGAGGGGCTTACTGTCGAAAGGCTGCGCAC 1140
QY 1141 CTGGGGGTGCGCTTACCGAGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
Db 1141 CTGGGGGTGCGCTTACCGAGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200

```

```

QY 1201 GCGGAGACACACGTCGGAGTAAGGAAGGATGATCGGCGG 1245
Db 1201 GCGGAGACACACGTCGGAGTAAGGAAGGATGATCGGCGG 1245

RESULT 3
LOCUS A87604 1260 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 1 from Patent WO9836089.
ACCESSION A87604
VERSION A87604.1 GI:6736244
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1260)
AUTHORS Flohe, L. and Singh, M.
TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL Patent: WO 9836089-A 20-AUG-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
source location/Qualifiers
1..1260
BASE COUNT 243 a 403 c 389 g 225 t
ORIGIN

Query Match 100.0%; Score 1245; DB 5; Length 1260;
Best Local Similarity 100.0%; Pred. No. 1.2e-171;
Matches 1245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTTGCAATTAATGCACTTTCTTACACTGAGGCTAGTACGAGGGGTAATC 60
Db 16 ATCTTGCAATTAATGCACTTTCTTACACTGAGGCTAGTACGAGGGGTAATC 75
QY 61 ATGCGCGTGGTATCC :CGAGACCAAAACAAGCAATTC VATTCCGGGTGSCATC 120
Db 76 ATGCGCGTGGTATCC :CGAGACCAAAACAAGCAATTC VATTCCGGGTGSCATC 135
QY 121 ACCCGCGCGCGCTGGGGAATTAACCCGTCGTGGCCATGAGTGTCTCATTCAGGAGGT 180
Db 136 ACCCGCGCGCGCTGGGGAATTAACCCGTCGTGGCCATGAGTGTCTCATTCAGGAGGT 195
QY 181 GCGGAGAGGGCTGCGCTATACCGAGCGGATTTCAAGCGGCGGAGCGCGCACTGATC 240
Db 196 GCGGAGAGGGCTGCGCTATACCGAGCGGATTTCAAGCGGCGGAGCGCGCACTGATC 255
QY 241 GGCACCGCGCGACAGGTCTGGGCGGACGCTGATTTATGCTCAAGGTCAAGAACCGATA 300
Db 256 GGCACCGCGCGACAGGTCTGGGCGGACGCTGATTTATGCTCAAGGTCAAGAACCGATA 315
QY 301 GCGGCGGAATACGGCGCGCTGCGACAGGGGAGATCTGTTCAAGGTCTTGCATTGGCC 360
Db 316 GCGGCGGAATACGGCGCGCTGCGACAGGGGAGATCTGTTCAAGGTCTTGCATTGGCC 375
QY 361 GCGTCACGTGCTTGCACCGATGCTGTGATTCGCGGACCAAGTCAATTTGCTTACGAG 420
Db 376 GCGTCACGTGCTTGCACCGATGCTGTGATTCGCGGACCAAGTCAATTTGCTTACGAG 435
QY 421 ACCGTCAGACCGCGGAGCGGCGCTACCGCTGCTTCCCGGATGAGCGAAGTCCGGGT 480
Db 436 ACCGTCAGACCGCGGAGCGGCGCTACCGCTGCTTCCCGGATGAGCGAAGTCCGGGT 495
QY 481 CGACTCGCGCGCGAGTTGGCGCTTACCACTGATGCGAACCAAGGGGGCGCGGTGTG 540
Db 496 CGACTCGCGCGCGAGTTGGCGCTTACCACTGATGCGAACCAAGGGGGCGCGGTGTG 555
QY 541 CTGATGGGCGGGGTGCCCGCGCTGGAACCGGCGAGCTGCTGATGATGCGGCGGCGCAC 600
Db 556 CTGATGGGCGGGGTGCCCGCGCTGGAACCGGCGAGCTGCTGATGATGCGGCGGCGACC 615

```

QY 601 GCCGCTACACGAGCCCGCATGCGCAAGGCGATGGGCGGACCGTTAGGTTCTAGAC 660
 |||||||
 Db 616 GCGCGCTACAAAGGAGCGCGATGCGCAACGCGCATGGGCGGACCGTTAGGTTCTAGAC 675
 |||||||
 QY 661 ATCAACATCGCAAACTTGGGCAACTGAGCGCGGAGTTCTGGCGCGGATCCACACTGCC 720
 |||||||
 Db 676 ATCAACATCGCAAACTTGGGCAACTGAGCGCGGAGTTCTGGCGCGGATCCACACTGCC 735
 |||||||
 QY 721 TACTCATCGGCTGAGCTGAGGAGTGGCGTCAAGGCGGACGCTGGTATGGGAGCC 780
 |||||||
 Db 736 TACTCATCGGCTGAGCTGAGGAGTGGCGTCAAGGCGGACGCTGGTATGGGAGCC 795
 |||||||
 QY 781 GTCTGTGTCAGGCGCGCAAGGCAACCAATTAGTCTGAAATTCATTGCTGGCATATG 840
 |||||||
 Db 796 GTCTGTGTCAGGCGCGCAAGGCAACCAATTAGTCTGAAATTCATTGCTGGCATATG 855
 |||||||
 QY 841 AAACAGAGTGGCTAGTGGTATATAGCATGACCAAGGCGGCTGTTTGAAGGCTCA 900
 |||||||
 Db 856 AAACAGAGTGGCTAGTGGTATATAGCATGACCAAGGCGGCTGTTTGAAGGCTCA 915
 |||||||
 QY 901 CGACCGACACCTACGACCAACCGGCTTCCGCTGACACAGCGTGTCTTACTGCGTG 960
 |||||||
 Db 916 CGACCGACACCTACGACCAACCGGCTTCCGCTGACACAGCGTGTCTTACTGCGTG 975
 |||||||
 QY 961 GCGAACATGCGCGCGCTGGTGGCGGAAGCGTACGCTACGCGGACCAAGCGAGATG 1020
 |||||||
 Db 976 GCGAACATGCGCGCGCTGGTGGCGGAAGCGTACGCTACGCGGACCAAGCGAGATG 1035
 |||||||
 QY 1021 CCGTATGTGCTGAGCTTGGCGGACCATGGCTGGGCGGCGGTCGCGTCAATCCGGA 1080
 |||||||
 Db 1036 CCGTATGTGCTGAGCTTGGCGGACCATGGCTGGGCGGCGGTCGCGTCAATCCGGA 1095
 |||||||
 QY 1081 CTACCCAAAGGTTCTTGGAGCAGCAAGAGGCGCTTACTGTCCGAACGGGTGCCACGAC 1140
 |||||||
 Db 1096 CTACCCAAAGGTTCTTGGAGCAGCAAGAGGCGCTTACTGTCCGAACGGGTGCCACGAC 1155
 |||||||
 QY 1141 CTGGGGGTGCGCTTACCGAGCGCGCGCGGCTGCTGACTCTGCGCGCTGTTAC 1200
 |||||||
 Db 1156 CTGGGGGTGCGCTTACCGAGCGCGCGCGGCTGCTGACTCTGCGCGCTGTTAC 1215
 |||||||
 QY 1201 GCGGACGACACGCTGGGAGTAAAGGAGCGATGATGTGCGCGCG 1245
 |||||||
 Db 1216 GCGGACGACACGCTGGGAGTAAAGGAGCGATGATGTGCGCGCG 1260
 |||||||

RESULT 4
 A89744
 LOCUS A89744 1260 bp DNA
 DEFINITION Sequence 1 from Patent WO9832862.
 ACCESSION A89744
 VERSION A89744.1 GI:6738278
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1260)
 AUTHORS Flohe, L. and Singh, M.
 TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
 JOURNAL Patent: WO 9832862-A 30-JUL-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source 1..1260
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 243 a 403 c 389 g 225 t
 ORIGIN
 Query Match 100.0%; Score 1245; DB 5; Length 1260;
 Best Local Similarity 100.0%; Pired. No. 1.2e-171;
 Matches 1245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTTGAGATTATGAACTTCTTACACTGAAAGCTACAGTATCGAGAGGGTAAATC 60

Db 16 ATCTTGAGATTATGAACTTCTTACACTGAAAGCTACAGTATCGAGAGGGTAAATC 75
 |||||||
 QY 61 ATGCGGTGGGTATTCGACCGGAGACCAAAACAGAAATTCATATCCGGGTGGCATTC 120
 |||||||
 Db 76 ATGCGGTGGGTATTCGACCGGAGACCAAAACAGAAATTCATATCCGGGTGGCATTC 135
 |||||||
 QY 121 ACCCGGCGGCTCGCGGAATTAACCCGTCGTCGAGAGGTGCTCATCCAGGAGGT 180
 |||||||
 Db 136 ACCCGGCGGCTCGCGGAATTAACCCGTCGTCGAGAGGTGCTCATCCAGGAGGT 195
 |||||||
 QY 181 GCGGAGAGGGCTCGGCTATACGACGCGGATTTCAAGGCGGAGCGCGCAACTGTTC 240
 |||||||
 Db 196 GCGGAGAGGGCTCGGCTATACGACGCGGATTTCAAGGCGGAGCGCGCAACTGTTC 255
 |||||||
 QY 241 GGCACCGCCGACAGGTGTGGGCGGCGGATTTTGTCTCAAGGTCAAGAACGATA 300
 |||||||
 Db 256 GGCACCGCCGACAGGTGTGGGCGGCGGATTTTGTCTCAAGGTCAAGAACGATA 315
 |||||||
 QY 301 GCGGCGGAATACGGCGGCTCGGACGAGGCGAGATCTGTTCAGTCTTTCATTTGGCC 360
 |||||||
 Db 316 GCGGCGGAATACGGCGGCTCGGACGAGGCGAGATCTGTTCAGTCTTTCATTTGGCC 375
 |||||||
 QY 361 GCGTACAGTGTTCGACCGGATGGGCTTGGGATTCCGCGACCAAGCTATGCTACGAG 420
 |||||||
 Db 376 GCGTACAGTGTTCGACCGGATGGGCTTGGGATTCCGCGACCAAGCTATGCTACGAG 435
 |||||||
 QY 421 ACCGTCCAGACCGCGGACGACCTACCCCTCTTCCCGGATGAGCAAGTCCCGGT 480
 |||||||
 Db 436 ACCGTCCAGACCGCGGACGACCTACCCCTCTTCCCGGATGAGCAAGTCCCGGT 495
 |||||||
 QY 481 CGACTCGCGCGGAGGTGGCGCTTACACCTATCGAAACCAAGGGGCGCGGTGG 540
 |||||||
 Db 496 CGACTCGCGCGGAGGTGGCGCTTACACCTATCGAAACCAAGGGGCGCGGTGG 555
 |||||||
 QY 541 CTGATGGGCGGAGGTGGCGCTTACACCTATCGAAACCAAGGGGCGCGGTGG 600
 |||||||
 Db 556 CTGATGGGCGGAGGTGGCGCTTACACCTATCGAAACCAAGGGGCGCGGTGG 615
 |||||||
 QY 601 GCGGCGTACACGACGCGCTTACCGCACGCGATGGGCGGACCGTTACGTTCTAGAC 660
 |||||||
 Db 616 GCGGCGTACACGACGCGCTTACCGCACGCGATGGGCGGACCGTTACGTTCTAGAC 675
 |||||||
 QY 661 ATCAACATCGCAAACTTGGGCAACTGAGCGCGGAGTTCTGGCGCGGATCCACACTGCC 720
 |||||||
 Db 676 ATCAACATCGCAAACTTGGGCAACTGAGCGCGGAGTTCTGGCGCGGATCCACACTGCC 735
 |||||||
 QY 721 TACTCATCGGCTGAGCTGAGGAGTGGCGTCAAGGCGGACGCTGGTATGGGAGCC 780
 |||||||
 Db 736 TACTCATCGGCTGAGCTGAGGAGTGGCGTCAAGGCGGACGCTGGTATGGGAGCC 795
 |||||||
 QY 781 GTCTGTGTCAGGCGCGCAAGGCAACCAATTAGTCTGAAATTCATTGCTGGCATATG 840
 |||||||
 Db 796 GTCTGTGTCAGGCGCGCAAGGCAACCAATTAGTCTGAAATTCATTGCTGGCATATG 855
 |||||||
 QY 841 AAACAGAGTGGCTAGTGGTATATAGCATGACCAAGGCGGCTGTTTGAAGGCTCA 900
 |||||||
 Db 856 AAACAGAGTGGCTAGTGGTATATAGCATGACCAAGGCGGCTGTTTGAAGGCTCA 915
 |||||||
 QY 901 CGACCGACACCTACGACCAACCGGCTTCCGCTGACACAGCGTGTCTTACTGCGTG 960
 |||||||
 Db 916 CGACCGACACCTACGACCAACCGGCTTCCGCTGACACAGCGTGTCTTACTGCGTG 975
 |||||||
 QY 961 GCGAACATGCGCGCGCTGGTGGCGGAAGCGTACGCTACGCGGTCGCAAGGAGATG 1020
 |||||||
 Db 976 GCGAACATGCGCGCGCTGGTGGCGGAAGCGTACGCTACGCGGTCGCAAGGAGATG 1035
 |||||||
 QY 1021 CCGTATGTGCTGAGCTTGGCGGACCATGGCTGGGCGGCGGTCGCGTCAATCCGGA 1080
 |||||||
 Db 1036 CCGTATGTGCTGAGCTTGGCGGACCATGGCTGGGCGGCGGTCGCGTCAATCCGGA 1095
 |||||||
 QY 1081 CTACCCAAAGGTTCTTGGAGCAGCAAGAGGCGCTTACTGTCCGAACGGGTGCCACGAC 1140
 |||||||

Db	1096	CTAACCAAAAGTCTTTTGAGAGCAAGAAAGGAGCGTACTGCTCCGAAGAGGATGCCACACGAC	1155
Qy	1141	CTGGGGGTGGCCGTTGACACGAGAGCCGCGACGCTGCTGGCCGACTCTCTCGGCGCGCTGTAC	1200
Dd	1156	CTGGGGGTGGCCGTTGACACGAGAGCCGCGACGCTGCTGGCTGACTCTCGGCGCGCTGTAC	1215
Qy	1201	GCCGAGCACACTGTGGAGTAAAGGAAAGCATATATCTCGGCGCGC	1245
Dd	1216	GCCGAGCACACTGTGGAGTAAAGGAAAGCATATATCTCGGCGCGC	1260

RESULT	5			
MTALADH				
LOCUS	2584 bp	DNA	BC ^T	01-DEC-1992
DEFINITION	M. tuberculosis gene for L-alanine dehydrogenase.			
ACCESSION	X63069	S36765		
VERSION	X63069.1	GI:44565		
KEYWORDS	alanine dehydrogenase; extracellular; intracellular.			
SOURCE	Mycobacterium tuberculosis.			
ORGANISM	Mycobacterium tuberculosis			

REFERENCE 1 (pages 1 to 2584)
AUTHORS Andersen, A.B.
TITLE Direct Submission
JOURNAL Submitted (08-NOV-1991) A.B. Andersen, Statens Seruminstitut

REFERENCE	AUTHORS	TITLE	JOURNAL
2 (Larsen J CO 2384*)	Andersen, A.B., Andersen, P. and Ljungqvist, L.	Structure and function of a 40,000-molecular-weight protein antigen of <i>Mycobacterium tuberculosis</i>	Infect. Immun. 60 (5), 2317-2323 (1992)

	Location/Qualifiers
FEATURES	

Source	Accession
1. 7236	/organism="Mycobacterium tuberculosis"
Strain-Erdman	/strain="Erdman"
Isolate-TMC 107 lot 9A-2"	/isolate="TMC 107 lot 9A-2"
db_xref="taxon:1773"	/db_xref="taxon:1773"
clone-11b-lambda gtl1"	/clone-11b-lambda gtl1"
clone-lambda Aa67"	/clone-lambda Aa67"
76. .1197	

BASE COUNT	463 a	871 c	819 g	431 t
ORIGIN				

Query Match	100.0%	Score 1245;	DB 1;	Length 2584;
Best Local Similarity	100.0%	Pred. No. 1,1e-171,		
Matches 1245;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

QY	121	ACC	CGGCGGGGCTCGCGGAACTAAACCGTGTGGCGAATGAGTGCATATCCAGCAAGT	180
Db	136	ACC	CGGCGGGGCTCGCGGAACTAAACCGTGTGGCGAATGAGTGCATATCCAGCAAGT	195
QY	181	GCG	GGAAGGGGCTCGGCTATCAACGCAGCGGGAATTTCAAGCGCGCAGGCGCAACTGGTC	240
Db	196	GCG	GGAAGGGGCTCGGCTATCAACGCAGCGGGAATTTCAAGCGCGCAGGCGCAACTGGTC	255
QY	241	GGC	ACCGCCACCAAGTGTGGGCCGACGCTGATTTATGTCTAAAGTAAAGAACGATA	300
Db	256	GGC	ACCGCCACCAAGTGTGGGCCGACGCTGATTTATGTCTAAAGTAAAGAACGATA	315
QY	301	GCG	GCGAATATACGGCGGCTCGGACACGGGCGAATCTTTCACGTTCTTGCAATTTGGC	360
Db	316	GCG	GCGAATATACGGCGGCTCGGACACGGGCGAATCTTTCACGTTCTTGCAATTTGGC	375
QY	361	GCG	CTACGTCTTCACACCGATGCGTGTGTTGATTCGCGCACCACTCAATTGCTACGAG	420
Db	376	GCG	CTACGTCTTCACACCGATGCGTGTGTTGATTCGCGCACCACTCAATTGCTACGAG	435
QY	421	ACG	CTCCAGACCGCGCGAGCGGCGACTAATCCCGATTGGCCGATGAGGAAGTGGCGGT	480
Db	436	ACG	CTCCAGACCGCGCGAGCGGCGACTAATCCCGATTGGCCGATGAGGAAGTGGCGGT	495
QY	481	CGA	CTCGCCCGCCAGGTTGGCGCTTACCACTGATGCGAATCCAAACCAAGGAGGCGCGGTG	540
Db	496	CGA	CTCGCCCGCCCGCAGGTTGGCGCTTACCACTGATGCGAATCCAAACCAAGGAGGCGCGGTG	555
QY	541	CTG	ATGGGGGGGGGTGCCCGGCTCGAATCCGGCGGAGAGTGTGTATGCGCGCGGACAC	600
Db	556	CTG	ATGGGGGGGGGGTGGCGGCTCGAATCCGGCGGAGAGTGTGTATGCGCGCGGACAC	615
QY	601	GCG	GCTATACAAGCGAGGCCCATCGGCAACGGCATGGGCGCACCGTTACGGTTCTAATC	660
Db	616	GCG	GCTATACAAGCGAGGCCCATCGGCAACGGCATGGGCGCACCGTTACGGTTCTAATC	675
QY	661	ATC	AACATCGAACAATTTGGGCAATCTGACGCCGAGTTCTGGCGCGGATCCACTCGC	720
Db	676	ATC	AACATATCGAACAATTTGGGCAATCTGACGCCGAGTTCTGGCGCGGATCCACTCGC	735
QY	721	TAC	ATATCGGGCTTACGAGCTCGAGGGTGGCGTAAACGTGCGACCTGGTATTTGGGGCC	780
Db	736	TAC	ATATATCGGGCTTACGAGCTCGAGGGTGGCGTAAACGTGCGACCTGGTATTTGGGGCC	795
QY	781	GTC	TGTGGTGCAGGGCGGCAAGGCCAACCCAATTAATGTCGAAATTACTGTGCGGCATATG	840
Db	796	GTC	TGTGGTGCAGGGCGGCAAGGCCAACCCAATTAATGTCGAAATTACTGTGCGGCATATG	855
QY	841	AAAC	CAGGTGGGACTGCTGTGATATATGCCATATGACACGAGGGGCGTGTTCGAAGCTCA	900
Db	856	AAAC	CAGGTGGGGGACTGCTGTGATATATGCCATATGACACGAGGGGCGTGTTCGAAGCTCA	915
QY	901	CGAC	GCGACACTTCGACACACCCGAGCTTGCGGCTGACAGACACGGTGTTTACTTGCCTG	960
Db	916	CGA	CGACGACACACTTCAGACACACCGAGCTTGCGGCTGACAGACACGGTGTTTACTTGCCTG	975
QY	961	GCG	AACATGCCCCGCTCTGGTGGCGAAGAGCTGACCTACGCGCTGACCAACGCGACGATG	1020
Db	976	GCG	AACATGCCCCGCTCTGGTGGCGAAGAGCTGACCTACGCGCTGACCAACGCGACGATG	1035
QY	1021	CCG	ATATGCTCGAGCTTGGCGACATATGGTGTGGCGGGCGCGCGCGTGCATATCCGGCA	1080
Db	1036	CCG	ATATGCTCGAGCTTGGCGACATATGGTGTGGCGGGCGCGCGCGTGCATATCCGGCA	1095
QY	1081	CTAC	GCAAAAGTCTTTTGAGAGCAGAAAGGGGCTTACTGTCCGAAGGGGTGGCCACCGAC	1140
Db	1096	CTAC	GCAAAAGGTTCTTTTGAGAGCAGAAAGGGGCTTACTGTCCGAAGGGGTGGCCACCGAC	1155
QY	1141	CTG	GGGGGTGCGTTACACGAGACCCGCGACGGTCTAGGCTGACTTCGGCGCTCTGTTAC	1200
Db	1156	CTG	GGGGGTGCGGCTTACACGAGACCCGCGACGGTCTAGGCTGACTTCGGCGCTCTGTTAC	1215
QY	1201	GCC	GACACACGTCGGGAGTAAAGGAACGATATGTGGCGCGC	1245


```

Db      1216  GCGGAGCACACGTCGGAGTAAGGAGGATGATCGCGCGC 1260
|||||
RESULT  6      MTW002      56414 bp      DNA      BCT      17-JUN-1998
LOCUS    MTW002      Mycobacterium tuberculosis H37Rv complete genome; segment 122/162.
DEFINITION
ACCESSION AL008967 AL123456
VERSION
KEYWORDS  AL008967.1 GI:3261491
SOURCE
ORGANISM  Mycobacterium tuberculosis.
          Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
          Actinomycetales; Corynebacterineae; Mycobacteriaceae;
          Mycobacterium.
REFERENCE 1 (bases 1 to 56414)
AUTHORS  Cole,S.T., Brosch,R., Parkhill,J., Garnier,T., Churcher,C.,
          Harris,D., Gordon,S.V., Eigmeier,K., Gas,S., Barry III,C.E.,
          Tekala,F., Badcock,K., Basham,D., Brown,D., Chillingworth,T.,
          Connor,R., Davies,R., Devlin,K., Felwell,T., Gentles,S.,
          Hamlin,N., Holroyd,S., Hornsby,T., Jagels,K., Krogh,A., McLean,J.,
          Moule,S., Murphy,L., Oliver,S., Osborn,J., Quail,M.A.,
          Rajandream,M.A., Rogers,R., Rutter,S., Seeger,K., Skelton,S.,
          Squares,S., Squires,R., Sulton,J.E., Taylor,K., Whitehead,S. and
          Barrett,J.G.
          Deciphering the biology of Mycobacterium tuberculosis from the
          complete genome sequence
          Nature 393 (6685), 537-544 (1998)
JOURNAL  Nature 393 (6685), 537-544 (1998)
MEDLINE  98295987
REMARK   Erratum: [[published erratum appears in Nature 1998 Nov
          12;396(6707):190]]
          2 (bases 1 to 56414)
          Parkhill,J.
          Direct Submission
          Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium
          tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome
          Trust Genome Campus, Hinxton, Cambridge, CB10 1SA Unite de Genetique
          Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,
          75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk
          On Jun 27, 1998 this sequence version replaced gi:2624256.
COMMENT  Notes:
          Details of M. tuberculosis sequencing at the Sanger Centre are
          available on the World Wide Web.
          (URL, http://www.sanger.ac.uk/Projects/M.tuberculosis/) CDS have
          been renumbered from the original cosmid submissions but the old
          gene designations are in brackets after the new gene numbers.
          Gene prediction was based on a Hidden Markov Model of TB genes
          implemented in tparse (Krogh) supplemented with visual inspection
          of positional base preference in codons, especially where there is
          an increase in the observed/expected third position G + C.
          CAUTION: In some cases we may not have predicted the correct
          initiation codon. Where possible we have chosen an initiation codon
          (atg, gtg, or ttg) which is preceded by an upstream ribosome
          binding site sequence (optimally 5-13bp before the initiation
          codon). If this cannot be identified we choose the most upstream
          initiation codon.
          Location/Qualifiers
            source
              1..56414
              /organism="Mycobacterium tuberculosis"
              /strain="H37Rv"
              /db_xref="taxon:1773"
              <1..556414
              /note="fragment designated v002. Does not represent a
              physical clone"
            misc_feature
              <1..233
              /organism="Mycobacterium tuberculosis"
              /strain="H37Rv"
              /db_xref="taxon:1773"
              /clone="Y154"
              /gene="recX"
              complement(3..527)
              complement(3..527)
            CDS
              gene
                complement(3..527)
              CDS
                complement(3..527)
          /gene="recX"
          /note="Rv2736c, (MTW002.01c), len: 174 aa, recx, similar
          to eg. RECX_PSEAE P37860 regulatory protein recx from
          Pseudomonas aeruginosa (155 aa), fasta scores: opt: 161
          z-score: 257.2 E(-): 3.e-07, 30.7% identity in 137 aa
          overlap. Overlaps and extends CDS from overlapping cosmid
          MTCY154.16c"
          /codon_start=1
          /product="recX"
          /protein_id="CA15532.1"
          /db_xref="GI:2624257"
          /db_xref="SPTREMBL:O33280"
          /translation="MTWSCPPSPTESEEDQARALCLRLTARSRAELAGQAKRGY
          PEDIGNRLDRLAIVGLVDTPDEKQVQSRPAARNAASRAELAEHAGVDDYIT
          VLGGIDAGSRGRARLREVRVIDIDGTDEARSRVLVLMARIGQLTACEV
          VLAELAEERRRRY"
          complement(493..2865)
          /gene="recA"
          complement(493..2865)
          /gene="recA"
          /note="Rv2737c, (MTW002.02c), len: 790 aa, recA,
          identical to recA_MYCTU P26345 recA protein (790 aa).
          contains self-splicing protein element (intein) from 2294
          to 969 (c), similar to intein II from TR:E332317
          (EMBL:Y13030) DNA-directed DNA polymerase (EC 2.7.7.7) from
          Thermococcus sp. (1829 aa), fasta scores: opt: 81 z-score:
          235.2 E(-): 6e-06, 24.6% identity in 183 aa overlap.
          Contains PS00017 ATP/GTP-binding site motif A (P-loop),
          PS00321 recA signature, and PS00861 protein splicing
          signature. See Davis et al, (1992) Cell 71(2):201-210"
          /codon_start=1
          /product="recA"
          /protein_id="CA15533.1"
          /db_xref="GI:2624258"
          /db_xref="SWISS-PROT:P26345"
          /translation="MTQTPREKALELAVAQIEKSYKSGVMRLGDEAROPISVPTG
          SIALDVALIGIGLPRGRVIEIGPESGKTTVAHVAANAAGVAALFIDAEHALDP
          DYARKLVDVDSLIVSOPDGEOLAEIADMLIRSGALDIYIVDSALVPRALDEEM
          GDSVGLQALMSQALSRKMTGALNNGTTFINOLBDRKIGVAPGSPETTGAKAF
          YASVRRRVERETKDGSTNAVGNTRKRVKKNCLAGCTIFPPTVGTTRIDVDYG
          RKLHVVAANKDGLHARPVASWFDQSTRTGLRAGIAIWAATDHWALVEYGRRA
          AGELRKGRDRAQPRFEDGDSAPIPDARHRLGLYLIGDRDQWAGKPPINEINOR
          ALIDVYRIATIGCAHPQGRISLAIAHPRGNGVADICQAGIYKLAEMKLTIPN
          WFEPPDIAADIVGNLGLFSPDGWVSREOTGALRVGTTSSQLAHQIMLLIRGV
          GSTYRDVDPKQRPSTVNGRIOSKQVFEVRSIGNMNTAPAEVPMGPRRAALIO
          AIPRATQRRRGSAITYLAAEMTDVANTYIDERGVTQAEAAATIGVAGSDPRGMQV
          LGASVLRDRVQAADALDLDKFLHDMIAEELRSVIREVLPTRARFTDLEVEELTL
          VAGSVVHNGSPPEKQAEFDILYKGISRGSGLIDMGVDGLIRKSGAMFTYEGEOLG
          QGKEMANFVLENADVDEIEKIKELKIGIGAVTDDPSNDGLVPAPVF"
          complement(790..2112)
          /gene="recA"
          /note="recA intein"
          complement(790..807)
          /gene="recA"
          /note="PS00861 protein splicing signature"
          complement(2197..2223)
          /gene="recA"
          /note="PS00321 recA signature"
          complement(2644..2667)
          /gene="recA"
          /note="PS00017 ATP/GTP-binding site motif A (P-loop)"
          complement(2873..2877)
          /note="possible RBS upstream of Rv2737c"
          complement(3247..3453)
          /gene="Rv2738c"
          complement(3247..3453)
          /gene="Rv2738c"
          /note="Rv2738c, (MTW002.03c), len: 68 aa, unknown.
          N-terminus is highly similar to the N-terminus of the
          upstream ORFMTW002.07c (78.4% identity in 37 aa overlap);
          also similar to AL020958|SC4H8_5 Streptomyces coelicolor

```

cosmid 4H8 (64aa) opt: 185 z-score: 283.5 E(): 2.9e-08;
39.7% identity in 63 aa overlap"

/codon_start=1
/transl_table=11
/product="hypothetical protein Rv2738c"
/protein_id="CA15534.1"
/db_xref="GI:2624260"
/db_xref="SPTREMBL:O33281"
/translation="MLAGVRLTEPHERVALHFGAAGSSVLLDHLVLPFGGRSAQAI
EDGVEPRDVRALCADFDVPHDRM"
/complement(3464..4630)
/gene="Rv2739c"
/note="Rv2739c. (MTV002.04c), len: 388 aa. Probable
transferase, similar to eg. TR-Q051560 (EMBL:L28170)
Pseudomonas aeruginosa thiamosyl transferase (426 aa),
fasta scores: opt: 178 z-score: 226.3 E(): 1.9e-05, 25.98
identity in 425 aa overlap. Equivalent to Mycobacterium
leprae protein MLCB33.02c (392 aa); fasta scores
opt: 244.723 MLCB33.2 Mycobacterium leprae cosmid B33 opt:
2112 z-score: 2364.5 E(): 0; 80.98 identity in 388 aa
overlap"

/codon_start=1
/transl_table=11
/product="hypothetical protein Rv2739c"
/protein_id="CA15535.1"
/db_xref="GI:2624261"
/db_xref="SPTREMBL:O33282"
/translation="MRVAVVAGPPDGHSPFALALCORRA..JPTFLTGYEMLEAAR
AAGIDAVDELDAATRDLDGARLHRAAQAQVAVNPRALPELVSIVTACGG
MAELLAGIPWELNPHLYLPKGLPPIGSLAIGTGRRLDATTMLALGRSMRG
LROKAAVEIGLPAKDPGRLRLIATLPALEVRPDPAAVAVVGRHPEPDRVA
IPAGGPPVAVAPSTALGTAGLEVALQSLTPGVSGSRLVSVRSGADLVPPM
AVAGISQAVELTRADIVICGGHGMVAKTLGAPMVVAVVGGDDQMLIARVYRQS
AVLIRPLADALVAANVNVLSPPRRRARARAAASVAGAAPVAVCHDALALAG"
4674..5123
/gene="Rv2740"
4674..5123
/gene="Rv2740"
/note="Rv2740. (MTV002.05), len: 149 aa. unknown, similar
to M. leprae hypothetical protein TR-Q049850 (EMBL:294723)
MLCB33.03 (B2235_F3_140) (178 aa), fasta scores: opt: 498
z-score: 538.9 E(): 7.3e-23, 51.6% identity in 161 aa
overlap"

/codon_start=1
/transl_table=11
/product="hypothetical protein Rv2740"
/protein_id="CA15536.1"
/db_xref="GI:2624262"
/db_xref="SPTREMBL:O33283"
/translation="MAELTESPPTETEARAVEAFNLQNEPDTVDALGDDL
VYENVFSIRIGRRTATLIRMOGRVFEVYIHIGADGAVALERDALIIPLRY
DFWVGVEVDGRTILWRYDYDMFKLLRGLVALVPSLKLTL"
5355..6932
/gene="pe_pgrs"
5355..6932
/gene="pe_pgrs"
/note="Rv2741. (MTV002.06), len: 525 aa; Member of M.
tuberculosis PE_pgrs subfamily, similar to many eg.

Query Match 98.1%; Score 1221.4; DB 1; Length 56414;
Best local Similarity 99.4%; Pred. No. 1.4e-168;
Matches 1238; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 1 ATCTTGACATTAATGACATTTCTTCACTAGACGTCATATCGAGAGGGGTAATC 60
DB 38201 ATCTTGACATTAATGACATTTCTTCACTAGACGTCATATCGAGAGGGGTAATC 38260

QY 61 ATGCGGCGTATTCGACGAGACCAAAACGAATTCATTCGCGGTGGCATC 120
DB 38261 ATGCGGCGTATTCGACGAGACCAAAACGAATTCATTCGCGGTGGCATC 38314

QY 121 ACCCGCGCGCGCGCGGGAAGTAACCCGTCGTGGCATGAGGTGCTCATCCAGGAGGT 180
DB 38315 ACCCGCGCGCGCGCGGGAAGTAACCCGTCGTGGCATGAGGTGCTCATCCAGGAGGT 38374

QY 181 GCGGAGAGGAGGCTCGGCTATCCAGCAGCGGATTTCAAGGCGGAGCGCGCAATGCTC 240
DB 38375 GCGGAGAGGAGGCTCGGCTATCCAGCAGCGGATTTCAAGGCGGAGCGCGCAATGCTC 38434

QY 241 GCGAGCGCGCGCGCGGAGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
DB 38435 GCGAGCGCGCGCGCGGAGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 38494

QY 301 GCGGCGGAAATACGCGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
DB 38495 GCGGCGGAAATACGCGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 38554

QY 361 GCGTCAGCTGCTTGCACCGATGCGGTTGATTCGCGGACGACGTCATTCGCTACGAG 420
DB 38555 GCGTCAGCTGCTTGCACCGATGCGGTTGATTCGCGGACGACGTCATTCGCTACGAG 38614

QY 421 ACCGTCAGACCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
DB 38615 ACCGTCAGACCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 38674

QY 481 CGACTCGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
DB 38675 CGACTCGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 38734

QY 541 CTGATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
DB 38735 CTGATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 38794

QY 601 GCGGCGTCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
DB 38795 GCGGCGTCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 38854

QY 661 ATCAACATGCAAACTTGGGCAACTGAGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
DB 38855 ATCAACATGCAAACTTGGGCAACTGAGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 38914

QY 721 TACGTCATGCGGCTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
DB 38915 TACGTCATGCGGCTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 38974

QY 781 GTCTGTGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
DB 38975 GTCTGTGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 39034

QY 841 AAACAGGTGCGGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 39035 AAACAGGTGCGGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 39094

QY 901 CGAGCGACGACCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 960
DB 39095 CGAGCGACGACCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 39154

QY 961 GCGAAGCATGCG 1020
DB 39155 GCGAAGCATGCG 39214

QY 1021 CCGTATGTCGTCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1080
DB 39215 CCGTATGTCGTCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 39274

QY 1081 CTAGCGCAAGGCTCTTGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1140
DB 39275 CTAGCGCAAGGCTCTTGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 39334

QY 1141 CTGGGCGTGCCTTGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1200
DB 39335 CTGGGCGTGCCTTGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 39394

QY 1201 GCCGAGCACACGTCGGGAGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1245

```

Db 39395 GCCGAGCAGACGCTGGGAGTAAGGGAAGCGATGATGTGCGCGCG 39439
|||||
RESULT 7
LOCUS AB7613 1236 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 10 from Patent WO9836089.
ACCESSION AB7613
VERSION AB7613.1 GI:6736253
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1236)
AUTHORS Flohe, L. and Singh, M.
TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL Patent: WO 9836089-A 20-AUG-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
source 1..1236
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 236 a 395 c 385 g 220 t
ORIGIN
Query Match 98.0%; Score 1220; DB 5; Length 1236;
Best Local Similarity 99.5%; Pred. No. 5.1e-168;
Matches 1236; Conservative 0; Mismatches 0; Indels 6; Gaps 1;
QY 1 ATCTTGAGATTATCGAACTTCTTCACACTGAGGTACAGTATGAGAGGGTAATC 60
Db 1 ATCTTGAGATTATCGAACTTCTTCACACTGAGGTATGAGATGAGAGGGTAATC 60
QY 61 ATGGCGCTGCTATTCGACCGAGACCAAAACAACGAATTCATTCGGGTGCCATC 120
Db 61 ATGGCGCTGCTATTCGACCGAGACCAAAACAACG-----AATTCGGGTGCCATC 114
QY 121 ACCCGGGCGGCGTGGGGAAGTAACCGGTGTGGCCATGAGTGTCTCATCGAGCAAGT 180
Db 115 ACCCGGGCGGCGTGGGGAAGTAACCGGTGTGGCCATGAGTGTCTCATCGAGCAAGT 174
QY 181 GCCGAGAGGGCTGGGCTATCACCGAGCGGATTTCAAGCGCGAGCGCAATGGTCTC 240
Db 175 GCCGAGAGGGCTGGGCTATCACCGAGCGGATTTCAAGCGCGAGCGCAATGGTCTC 234
QY 241 GGCACCGCCGACAGGTGTGGCGGAGCGCTATTATTGCTCAAGGTCAAAAGAACCGATA 300
Db 235 GGCACCGCCGACAGGTGTGGCGGAGCGCTATTATTGCTCAAGGTCAAAAGAACCGATA 294
QY 301 GCGCGGGAATACGGCGCGCTCGACAGCGGAGATCTTGTTCAGCTTCTTCATTGGCC 360
Db 295 GCGCGGGAATACGGCGCGCTCGACAGCGGAGATCTTGTTCAGCTTCTTCATTGGCC 354
QY 361 GCGTACAGTCTTCACCGATGCGTGTGTGATTCGGGACACAGTCAATTCCTTACGAG 420
Db 355 GCGTACAGTCTTCACCGATGCGTGTGTGATTCGGGACACAGTCAATTCCTTACGAG 414
QY 421 ACCGTCCAGACCGCGGAGCGGCACTACCTCTGCTTGGCCCCGATGAGCGAAGTCCGGGT 480
Db 415 ACCGTCCAGACCGCGGAGCGGCACTACCTCTGCTTGGCCCCGATGAGCGAAGTCCGGGT 474
QY 481 CGACTCGCGCGCGAGGTGGGCTTACCACTGATGGAAACCAAGGGGGCGCGGGTGTG 540
Db 475 CGACTCGCGCGCGAGGTGGGCTTACCACTGATGGAAACCAAGGGGGCGCGGGTGTG 534
QY 541 CTGATGGCGGGGCTGCCGCGGTGAACCGGCGACGTCGTGTGATCGCGCGCGGACCC 600
Db 535 CTGATGGCGGGGCTGCCGCGGTGAACCGGCGACGTCGTGTGATCGCGCGCGGACCC 594
QY 601 GCGGCGTCAACGACGCCCGCATGCCAACGCGATGGCAGCGACCGTTACGCTTACAC 660
|||||

```

```

Db 595 GCCGCGTACACGACGCCCGCATCGCCAAAGCGATGGGGCGGACCTTACGTTACAC 654
QY 661 ATCAACATGACAAACTTGGCACTGACGCGCGAGTCTGGGCGGATTCACACTGCG 720
Db 655 ATCAACATGACAAACTTGGCACTGACGCGCGAGTCTGGGCGGATTCACACTGCG 714
QY 721 TACTCATCGGCTTACGAGCTCGAGGGTGGCGTCAAAAGTCCGCACTGATGGGGCC 780
Db 715 TACTCATCGGCTTACGAGCTCGAGGGTGGCGTCAAAAGTCCGCACTGATGGGGCC 774
QY 781 GTCCTGTCGACGCGCCCAAGGCAACCAATTAGTCTCAATTCATTTGCGCATATG 840
Db 775 GTCCTGTCGACGCGCCCAAGGCAACCAATTAGTCTCAATTCATTTGCGCATATG 834
QY 841 AAACCAAGTCCGCTACTGCTGATATAGCCATCGACCGAGGGGCGCTGTTCCAGGCTCA 900
Db 835 AAACCAAGTCCGCTACTGCTGATATAGCCATCGACCGAGGGGCGCTGTTCCAGGCTCA 894
QY 901 CGACCGACACCTACGACCCACCGGAGCTTGGCGGTGACAGACAGCGTGTTCATGCTG 960
Db 895 CGACCGACACCTACGACCCACCGGAGCTTGGCGGTGACAGACAGCGTGTTCATGCTG 954
QY 961 GCGAAGATGCGCGCTCGGTGCGGAAGAGTGCACCTACGCGCTGACCAACGAGGATG 1020
Db 955 GCGAAGATGCGCGCTCGGTGCGGAAGAGTGCACCTACGCGCTGACCAACGAGGATG 1014
QY 1021 CCGTATGTGCTCAGCTTGGCCGACATGGCTGGCGGGCGGCTGGCGGTGCAATCGGCA 1080
Db 1015 CCGTATGTGCTCAGCTTGGCCGACATGGCTGGCGGGCGGCTGGCGGTGCAATCGGCA 1074
QY 1081 CTAGCCAAAGTCTTTCGACGACGCAAGAGGCGCTTCTCTCGAAGGGGTGGCCACGAC 1140
Db 1075 CTAGCCAAAGTCTTTCGACGACGCAAGAGGCGCTTCTCTCGAAGGGGTGGCCACGAC 1134
QY 1141 CTGGGGGTGCGCTTACGACGAGCCCGCAGGTGCTGGCTGACTCTGCGCGCTGTTAC 1200
Db 1135 CTGGGGGTGCGCTTACGACGAGCCCGCAGGTGCTGGCTGACTCTGCGCGCTGTTAC 1194
QY 1201 GCGGAGCAGCGTGGGAGTAAGGGAAGCATGATCTGGCC 1242
Db 1195 GCGGAGCAGCGTGGGAGTAAGGGAAGCATGATGATCTGGCC 1236
|||||
RESULT 8
LOCUS AB9753 1236 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 10 from Patent WO9832862.
ACCESSION AB9753
VERSION AB9753.1 GI:6748287
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1236)
AUTHORS Flohe, L. and Singh, M.
TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
JOURNAL Patent: WO 9832862-A 30-JUL-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
source 1..1236
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 236 a 395 c 385 g 220 t
ORIGIN
Query Match 98.0%; Score 1220; DB 5; Length 1236;
Best Local Similarity 99.5%; Pred. No. 5.1e-168;
Matches 1236; Conservative 0; Mismatches 0; Indels 6; Gaps 1;
QY 1 ATCTTGAGATTATCGAACTTCTTCACACTGAGGTACAGTATGAGAGGGTAATC 60
Db 1 ATCTTGAGATTATCGAACTTCTTCACACTGAGGTATGAGATGAGAGGGTAATC 60

```

OY	61	ATGGCGCTGGTATTCCGACCGAGACCCAAAAACAAGAAATTCAAATTCGCGGGTGGCCATC	120
Db	61	ATGGCGCTGGTATTCCGACCGAGACCCAAAAACAAG-----AAATTCGCGGGTGGCCATC	114
OY	121	ACCCGCGCGCGCTGCGGGAACCTAACCCGTCGTGGCCATGAGTGCTCATCCAGCAGGT	180
Db	115	ACCCGCGCGCGCTGCGGGAACCTAACCCGTCGTGGCCATGAGTGCTCATCCAGCAGGT	174
OY	181	GCCGAGAGGGCTGGCTATCACCGACCGGATTTCAAGCGGCAGGCCCACTGGTC	240
Db	175	GCCGAGAGGGCTGGCTATCACCGACCGGATTTCAAGCGGCAGGCCCACTGGTC	234
OY	241	GGCACCGCGCGACAGAGTGTGGGGCGAGCGCTGATTTATTTGCTCAAGSTCAAAAGCAGATA	300
Db	235	GGCACCGCGCGACAGAGTGTGGGGCGAGCGCTGATTTATTTGCTCAAGSTCAAAAGCAGATA	294
OY	301	GCGCGGAATACGGCGCGCTGCGACACGGGCAAGTCTTTTCACGTTCCTTGCAATTTGGCC	360
Db	295	GCGCGGAATACGGCGCGCTGCGACAGGGGCAAGTCTTTTCACGTTCCTTGCAATTTGGCC	354
OY	361	GCGTACGTGTTGACCCGATAGCGTGTGGATTTCCGGC-ACGTCANATGGCT-CAAG	420
Db	355	GCGTACGTGTTGACCCGATAGCGTGTGGATTTCCGGCAGCACAGTCATAGCTTACAG	414
OY	421	ACCGCTCAGACCGCGCGACGGGCGCACTACCCCTGCTTCCCGCATGTGACCAAGTCGCGGT	480
Db	415	ACCGCTCAGACCGCGCGAGAGGGGCACTACCCCTGTTCCCGCATGTAGCAAGTCGCGGT	474
OY	481	CGACTCGCGCGCCCAAGTGTGGCGCTTACCACCTGATGGAACCCAGGGGCGCGGTGTG	540
Db	475	CGACTCGCGCGCCCAAGTGTGGCGCTTACCACCTGATGGAACCCAGGGGCGCGGTGTG	534
OY	541	CTGATGGGCGGGGAGCCCGGGGCTGGAACCGGCGACGCTGTGGTGAATCGCGCGGACAC	600
Db	535	CTGATGGGCGGGGAGTCCCGGGGCTGGAACCGGCGACGCTGTGGTGAATCGCGCGGACAC	594
OY	601	GCCGCTACAAACGACGCCCGCATCGCCAAAGGCAAGGGCGCGACCGTTACGTTAGAC	660
Db	595	GCCGCTACAAACGACGCCCGCATCGCCAAAGGCAAGGGCGCGACCGTTACGTTACAGAC	654
OY	661	ATCAACATCGCAAACTTTGGGCMACTGACAGCGGAGTTTGGCGCGGATTCACAACTGCG	720
Db	655	ATCAACATCGCAAACTTTGGGCMACTGACAGCGGAGTTTGGCGCGGATTCACAACTGCGC	714
OY	721	TACTCATGCGCTACGAGAGCTGAGAGGTGCGCGTCAAAAGTSCCGACCTGGTGAATGGGCGC	780
Db	715	TACTCATGCGCTACGAGAGCTGAGAGGTGCGCGTCAAAAGTSCCGACCTGGTGAATGGGCGC	774
OY	781	GTCCTGTGTCAGGCGCCAAAGCAACCCAAATTTAGTCTCGAATTCACCTTGTGCGCATATG	840
Db	775	GTCCTGTGTCAGGCGCCAAAGCAACCCAAATTTAGTCTCGAATTCACCTTGTGCGCATATG	834
OY	841	AAACAGGTGTGGATACGTGGGAAATATAGCATGACAGAGGGGCGGTTTCGAAGGCTCA	900
Db	835	AAACAGGTGTGGATACGTGGGAAATATAGCATGACAGAGGGGCGGTTTCGAAGGCTCA	894
OY	901	CGACCGACCACTACGACACCAACCGAGCTTGCCGTGACGACACGCTGTTTACTGCGTG	960
Db	895	CGACCGACCACTACGACACCAACCGACCTTGCCGTGACGACACGCTGTTTACTGCGTG	954
OY	961	GCGAACAATGCGCGCTGCGGTGGCGAABAGTTCGACCTAGCGCGGTGACCAAGCGAGATG	1020
Db	955	GCGAACAATGCGCGCTGCGGTGGCGAABAGTTCGACCTAGCGCGGTGACCAAGCGAGATG	1014
OY	1021	CCGATGTGTCGAGCTTGCCGACCAATGCTGGCGGGCGCGCGTGCCTGTCGATCCGGCA	1080
Db	1015	CCGATGTGTCGAGCTTGCCGACCAATGCTGGCGGGCGCGCGTGCCTGTCGATCCGGCA	1074
OY	1081	CTAGCCAAAGGTCTTTTGAGCGAAGAGGGGCGTTACTGTCCGAAAGGGTGGCCACAGAC	1144
Db	1075	CTAGCCAAAGGTCTTTTGAGCGAAGAGGGGCGTTACTGTCCGAAAGGGTGGCCACAGAC	1138

QY	1141	CGGGGGGTCGGCTTACCGAGCCCGCCAGCGTGTGCGCTACTGCTCTGGGCCCTGTTAC	1200
DB	1135	CTGGGGGTCGGCTTACCGAGCCCGCCAGCGTGTGCGCTACTGCTCTGGGCCCTGTTAC	1194
QY	1201	GCCGAGCACACGTCGGAGTAGGAGGACGATGATGTCGCC	1242
DB	1195	GCCGAGCACACCTCGGGAGTAAAGGAGACGATGATGTCGCC	1236

RESULT	9
LOCUS	A87607
DEFINITION	Sequence 4 from Patent WO9836089.
ACCESSION	A87607
VERSION	A87607.1
KEYWORDS	GI:6736247
SOURCE	unidentified.
ORGANISM	unidentified.
REFERENCE	1 (bases 1 to 1237)
AUTHORS	Flohe, L. and Singh, M.
TITLE	TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL	Patent: WO 9836089-A 20-AUG-1998;
FEATURES	FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
source	Location/Qualifiers
BASE COUNT	236 a 394 c 386 g 221 t
ORIGIN	1..1237

Query Match	97.9%	Score 1219.4	DB 5	Length 1237
Best Local Similarity	99.4%	Pred. No. 6.3e-168		
Matches 1236	Conservative	0	Mismatches 1	Indels 6
			Gaps	

QY	1	ATCTTCGACATTATTCGAACTTTCTTCACACTGAAGCGTACGATTCGAGAGGGTAATC	60
DB	1	ATCTTCGACAGTTAAATCGAACTTTCTTCATCTAGGAAGCGTACGATTCGAGAGGGTAATC	60
QY	61	ATGCGCGTCGCGATTCGCGCGAGACCAAAACAAAGAAATTCATTCGGGTGGCCATC	120
DB	61	ATGCGCGTCGCGATTCGCGCGAGACCAAAACAAAGAAATTCATTCGGGTGGCCATC	114
QY	121	ACCCGCGCGCGCGTTCGCGCAACTAACCCGCTCGTGCCATGAGGTGCTGAAACCGACAGGT	180
DB	115	ACCCGCGCGCGCGTTCGCGCAACTAACCCGCTCGTGCCATGAGGTGCTGATCCAGCGAGGT	174
QY	181	GCCGAGAGAGGCTCGCGCTATTCACGACGCGGATTTCAAGGGCGGACGCGCACTGGTC	240
DB	175	GCCGAGAGAGGCTCGCGCTATTCACGACGCGGATTTCAAGGGCGGACGCGCACTGGTC	234
QY	241	GGCACCGCGGACGAGGTGTGGGGCGGACCGCTGATTATGTCAAGTCAAAAGCGATA	300
DB	235	GGCACCGCGGACGAGGTGTGGGGCGGACCGCTGATTATGTCTAAAGTCAAAAGCGATA	294
QY	301	GCGGCGGAATACGCGCGCTCGGACACGCGGACGATCTGTTCAGTTCCTTGCAATTGGCC	360
DB	295	GCGGCGGAATACGCGCGCTCGGACACGCGGACGATCTGTTCAGTTCCTTGCAATTGGCC	354
QY	361	GCGTACGCTGCTTCGACCGGATGCTGTGTGGATTCGCGGACCGACGATCAATTGCTACGAG	420
DB	355	GCGTACGCTGCTTCGACCGGATGCTGTGTGGATTCGCGGACCGACGATCAATTGCTACGAG	414
QY	421	ACCGTTCAGACCGCGGACGCGGCGCTATACCCGCTGCGCCGATAGGAAGAAATGCCCGGT	480
DB	415	ACCGTTCAGACCGCGGACGCGGCGCTATACCCGCTGCTGCGCCGATAGGAAGAAATGCCCGGT	474
QY	481	CGACTTCGCGCCGACGATTTGGGCTTACCACTGATGCGAACCAGAGGGGCGCGCGTGTG	540
DB	475	CGACTTCGCGCCGACGATTTGGGCTTACCACTGATGCGAACCAGAGGGGCGCGCGTGTG	534
QY	541	CTGATGGGCGGGGTCGCCGCGCTCGAACCGCGCGAGCTGTGATGATCGGCGCGGACCG	600

```

|||||
Db 535 CTGATGGGGGGGGTGGCCGGCTCCAGACCGGCGAGCGTGTGTATGGCGCGCCGACCC 594
|||
Qy 601 GCGGGCTACAGCAGCAGCCGCATCGCCAAAGCATGGCGCGCAGCGTTACGTTTACAC 660
|||
Db 595 GCGGGCTACAGCAGCAGCCGCATCGCCAAAGCATGGCGCGCAGCGTTTACGTTTACAC 654
|||
Qy 661 ATCAACATGACAAATCTGGGCACTCGAGCGCGGATTTGGCGCGGATTCACACTGCG 720
|||
Db 655 ATCAACATGACAAATCTGGGCACTCGAGCGCGGATTTGGCGCGGATTCACACTGCG 714
|||
Qy 721 TACTCATCGCGCTACGAGCTCGAGGGTGGCTTAAAGCTGCGGACCTGATTTGGGCGC 780
|||
Db 715 TACTCATCGCGCTACGAGCTCGAGGGTGGCTTAAAGCTGCGGACCTGATTTGGGCGC 774
|||
Qy 781 GTCTGTGTCGAGCGCGCGCAAGGCGCAATTAATGTCGAATTCATTTGCGCGCATATG 840
|||
Db 775 GTCTGTGTCGAGCGCGCGCAAGGCGCAATTAATGTCGAATTCATTTGCGCGCATATG 834
|||
Qy 841 AAACGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
|||
Db 835 AAACGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 894
|||
Qy 901 CGACGACCACTACGACCACTACGACCACTACGACCACTACGACCACTACGACCACTACG 960
|||
Db 895 CGACGACCACTACGACCACTACGACCACTACGACCACTACGACCACTACGACCACTACG 954
|||
Qy 961 GCAACATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
|||
Db 955 GCAACATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1014
|||
Qy 1021 CGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
|||
Db 1015 CGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1074
|||
Qy 1081 CTAGCCAAAGTCTTTCGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1140
|||
Db 1075 CTAGCCAAAGTCTTTCGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1134
|||
Qy 1141 CTGGGGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
|||
Db 1135 CTGGGGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1194
|||
Qy 1201 GCGGAGCAGCGTGGAGTGAAGGAAAGGATGTCGCGCG 1243
|||
Db 1195 GCGGAGCAGCGTGGAGTGAAGGAAAGGATGTCGCGCG 1237
|||

```

```

RESULT 10
LOCUS A89747 1237 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 4 from Patent WO9832862.
ACCESSION A89747
VERSION A89747.1 GI:6738281
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1237)
AUTHORS Flohe, L. and Singh, M.
TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
JOURNAL Patent: WO 9832862-A 30-JUL-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
Location/Qualifiers
1..1237
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 236 a 394 c 386 g 221 t
ORIGIN
Query Match 97.98; Score 1219.4; E 5; Length 1237;
Best Local Similarity 99.48; Pred. No. 6.3e-108;

```

```

Matches 1236; Conservative 0; Mismatches 1; Indels 6; Gaps 1;
Qy 1 ATCTTGAGATTAATGACATTTCTTACACTGAGGCTACATATGAGAGGGTAATC 60
|||
Db 1 ATCTTGAGATTAATGACATTTCTTACACTGAGGCTACATATGAGAGGGTAATC 60
|||
Qy 61 ATGCGCTGCGTATTCGCGAGAGAGCAAAACAGCAATTCGCGGTGCGCATC 120
|||
Db 61 ATGCGCTGCGTATTCGCGAGAGAGCAAAACAGCAATTCGCGGTGCGCATC 114
|||
Qy 121 ACCCGGCGCGGCTGCGGAGAACTAACCGCTGCGGCTGAGGCTGATTCAGGAGGT 180
|||
Db 115 ACCCGGCGCGGCTGCGGAGAACTAACCGCTGCGGCTGAGGCTGATTCAGGAGGT 174
|||
Qy 181 GCGGAGAGG 3CTGCGCTATCACCGAGCGGATTTCAAGGCGGCGGCAACTGTC 240
|||
Db 175 GCGGAGAGGCTCGGCTA. CACGAGCGGATTTCAAGGCGGCGGCAACTGTC 234
|||
Qy 241 GCGACCGCGGACAGGCTGCGGAGAGGCTGATTTATGCTCAAGGCTCAAGAGAGGAT 300
|||
Db 235 GCGACCGCGGACAGGCTGCGGAGAGGCTGATTTATGCTCAAGGCTCAAGAGAGGAT 294
|||
Qy 301 GCGGCGGATACGCGCGCTGCGAGAGGAGATCTTGTTCAGGCTTTCGATTTGGCC 360
|||
Db 295 GCGGCGGATACGCGCGCTGCGAGAGGAGATCTTGTTCAGGCTTTCGATTTGGCC 354
|||
Qy 361 GCGTACGCTGCTGACCGATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
|||
Db 355 GCGTACGCTGCTGACCGATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 414
|||
Qy 421 ACCGTCAGAGCGCGCGAGGCGGAGCTACCGCTGCTGCGGATGAGCGAAGTGGCGGT 480
|||
Db 415 ACCGTCAGAGCGCGCGAGGCGGAGCTACCGCTGCTGCGGATGAGCGAAGTGGCGGT 474
|||
Qy 481 CGACTCGCGCGCCAGGTTGGCTTACCACTGATGCAAGGAGGCGCGGCTG 540
|||
Db 475 CGACTCGCGCGCCAGGTTGGCTTACCACTGATGCAAGGAGGCGCGGCTG 534
|||
Qy 541 CTGATGGCGGGGGTCCCGGGGTGGAACCGGCGAGCGTGGTGGATCGGCGCGGAGC 600
|||
Db 535 CTGATGGCGGGGGTCCCGGGGTGGAACCGGCGAGCGTGGTGGATCGGCGCGGAGC 594
|||
Qy 601 GCGGCTCAACAGCAGCGCTGATGCGCAACGAGATGGCGGCGGACGTTACGTTCTAGAC 660
|||
Db 595 GCGGCTCAACAGCAGCGCTGATGCGCAACGAGATGGCGGCGGACGTTACGTTCTAGAC 654
|||
Qy 661 ATCAACATGACAAATCTGCGTACGAGCGCGGATTCGCGGCGGATCAACACTGCG 720
|||
Db 655 ATCAACATGACAAATCTGCGTACGAGCGCGGATTCGCGGCGGATCAACACTGCG 714
|||
Qy 721 TACTCATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
|||
Db 715 TACTCATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 774
|||
Qy 781 GTCTGTGTCGAGCGCGCGCAAGGCGCAATTAATGTCGAATTCATTTGCGCGCATATG 840
|||
Db 775 GTCTGTGTCGAGCGCGCGCAAGGCGCAATTAATGTCGAATTCATTTGCGCGCATATG 834
|||
Qy 841 AAACGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
|||
Db 835 AAACGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 894
|||
Qy 901 CGACGACCACTACGACCACTACGACCACTACGACCACTACGACCACTACGACCACTACG 960
|||
Db 895 CGACGACCACTACGACCACTACGACCACTACGACCACTACGACCACTACGACCACTACG 954
|||
Qy 961 GCAACATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
|||
Db 955 GCAACATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1014
|||
Qy 1021 CGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
|||
Db 1015 CGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1074
|||

```

QY 1081 CTAGCCAAAGTCTTTGACGACGAGAGGGCGTTACTGTCGGAAGGGGTGGCCACCGAC 1140
 |||||
 Db 1075 CTAGCCAAAGTCTTTGACGACGAGAGGGCGTTACTGTCGGAAGGGGTGGCCACCGAC 1134
 |||||
 QY 1141 CTAGGGGTGCGCTTCACGAGCCCGCGAGCGTCTGGCTGACTCTTGCGCGCTGTTAC 1200
 |||||
 Db 1135 CTAGGGGTGCGCTTCACGAGCCCGCGAGCGTCTGGCTGACTCTTGCGCGCTGTTAC 1194
 |||||
 QY 1201 GCGGAGCACACGTCGGAGTAAGGAAGCGATGATGCGCGC 1243
 |||||
 Db 1195 GCGGAGCACACGTCGGAGTAAGGAAGCGATGATGCGCGC 1237
 |||||
 RESULT 11
 A87606 1235 bp DNA PAT 22-JAN-2000
 LOCUS A87606
 DEFINITION Sequence 3 from Patent WO9836089.
 ACCESSION A87606
 VERSION A87606.1 GI:6736246
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1235)
 AUTHORS Flohe, L. and Singh, M.
 TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
 JOURNAL Patent: WO 9836089-A 20-AUG-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source /organism="unidentified"
 1. .1235
 /db_xref="taxon:32644"
 BASE COUNT 236 a 394 c 382 g 220 t 3 others
 ORIGIN
 Query Match 97.7% Score 1216; DB 5; Length 1235;
 Best Local Similarity 99.3% Pred. No. 1.9e-167;
 Matches 123; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

QY 481 CGACTGCGCGCCAGGTTGGCGTTACCACTGATGCGAACCAGGGGGCGGCGTGTG 540
 |||||
 Db 475 CGACTGCGCGCCAGGTTGGCGTTACCACTGATGCGAACCAGGGGGCGGCGTGTG 534
 |||||
 QY 541 CTGATGGGGGGGTGCGCGCGCTGCAACCGCGCGAGCTGTGTGATCGGCGCGCGAC 600
 |||||
 Db 535 CTGATGGGGGGGTGCGCGCGCTGCAACCGCGCGAGCTGTGTGATCGGCGCGCGAC 594
 |||||
 QY 601 GCGGCTACAGCAGCAGCCCGCATCGGCAACGCGATGGGGCGGCGCTTACGATGAC 660
 |||||
 Db 595 GCGGCTACAGCAGCAGCCCGCATCGGCAACGCGATGGGGCGGCGCTTACGATGAC 654
 |||||
 QY 661 ATCAATGACACAACTTGGCACTGACGCGGAGTGTGCGCGGATTCACACTGCG 720
 |||||
 Db 655 ATCAATGACACAACTTGGCACTGACGCGGAGTGTGCGCGGATTCACACTGCG 714
 |||||
 QY 721 TACTCATCGGCTTACGAGCTCGAGGTCGCCCTGAAACGTGCGCACTGTGATGGGCG 780
 |||||
 Db 715 TACTCATCGGCTTACGAGCTCGAGGTCGCCCTGAAACGTGCGCACTGTGATGGGCG 774
 |||||
 QY 781 GTCTGTGCGGCGCGCAAGGCAACCAATTGCTCGAATTCAGTTGCGGCGCATATG 840
 |||||
 Db 775 GTCTGTGCGGCGCGCAAGGCAACCAATTGCTCGAATTCAGTTGCGGCGCATATG 834
 |||||
 QY 841 AACCAAGTGCAGTGTGATAGCATAGCAGCAGGCGGCTGTTGCAAGGCTCA 900
 |||||
 Db 835 AACCAAGTGCAGTGTGATAGCATAGCAGCAGGCGGCTGTTGCAAGGCTCA 894
 |||||
 QY 901 CGACCGACACCTTACGACACCGCGAGCTTCCGCTGACACACGCTGTTTACTCGTG 960
 |||||
 Db 895 CGACCGACACCTTACGACACCGCGAGCTTCCGCTGACACACGCTGTTTACTCGTG 954
 |||||
 QY 961 GCGAACATGCGCGCGCTCGGTGCGGAGAGCTGCACTAGCGGCTGACCAAGCGAGATG 1020
 |||||
 Db 955 GCGAACATGCGCGCGCTCGGTGCGGAGAGCTGCACTAGCGGCTGACCAAGCGAGATG 1014
 |||||
 QY 1021 CCGTATGTGCTGAGCTTGCAGCAATGCTGCGGCGGCGCTGCGGTGCAATCCGCA 1080
 |||||
 Db 1015 CCGTATGTGCTGAGCTTGCAGCAATGCTGCGGCGGCGCTGCGGTGCAATCCGCA 1074
 |||||
 QY 1081 CTAGCCAAAGTCTTTGACGACGAGAGGGGCTTACTGTCGAAGGGGTGGCCACCGAC 1140
 |||||
 Db 1075 CTAGCCAAAGTCTTTGACGACGAGAGGGGCTTACTGTCGAAGGGGTGGCCACCGAC 1134
 |||||
 QY 1141 CTGGGGTGGCGTTACCGAGCCCGCGAGCGTCTGGCGTCACTGCGGCGCTGTTAC 1200
 |||||
 Db 1135 CTGGGGTGGCGTTACCGAGCCCGCGAGCGTCTGGCGTCACTGCGGCGCTGTTAC 1194
 |||||
 QY 1201 GCGGAGCACACGTCGGAGTAAGGAAGCGATGATGTCGCG 1241
 |||||
 Db 1195 GCGGAGCACACGTCGGAGTAAGGAAGCGATGATGTCGCG 1235
 |||||
 RESULT 12
 A89746 1235 bp DNA PAT 22-JAN-2000
 LOCUS A89746
 DEFINITION Sequence 3 from Patent WO9832862.
 ACCESSION A89746
 VERSION A89746.1 GI:6738280
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1235)
 AUTHORS Flohe, L. and Singh, M.
 TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
 JOURNAL Patent: WO 9832862-A 30-JUL-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source /organism="unidentified"
 1. .1235
 /db_xref="taxon:32644"
 BASE COUNT 236 a 394 c 382 g 220 t 3 others

ORIGIN

Query Match	97.7%	Score 1216;	DB 5;	Length 1235;
Best Local Similarity	99.3%	Pred. No. 1.9e-167;		
Matches 1232; Conservative	0;	Mismatches 3;	Indels 6;	Gaps 1

QY	1	ATCTGGAGATTATTCGAACCTTTCTTCACACTCAACGCTACAGTATCTCGAGGGGTATTC	60
Db	1	ATCTTCGACGATTATTCGAACCTTTCTTCACACTCAACGCTACAGTATCTCGAGGGGTATTC	60
QY	61	ATGCGCGTGGTATTTCCGACCGAGACCCAAAACACGAATTCCATTCCGGGTGGCATTC	120
Db	61	ATGCGCGTGGTATTTCCGACCGAGACCCAAAACACG-----AATTCCGGGTGGCCATTC	114
QY	121	ACCCCGGGCCGCGTGGGGAACCTAACCCGCTGGGCGATGAGGTGCTCATCCAGGCAGGT	180
Db	115	ACCCCGGGCCGCGTGGGGAACCTAACCCGCTGGGCGATGAGGTGCTCATCCAGGCAGGT	174
QY	181	GCCGGAGAGGGCTGCGCTATCCACGACGGGATTTCAAGGGCGGACGGCCGCACACTGGTC	240
Db	175	GCCGGAGAGGGCTGCGCTATCCACGACGGGCGGATTTCAAGGGCGGACGGCCGCACACTGGTC	234
QY	241	GGCACCGCGCCGACGAGGTGTGGGGCGAGCGGTATTTCTCTCAAGSTCAAAABAACCGATA	300
Db	235	GGCACCGCGCCGACGAGGTGTGGGGCGAGCGGTATTTCTCTCAAGSTCAAAABAACCGATA	294
QY	301	GCGGCGGAAATACGGCCGCCCTTCGCAACAGGGGACAGATCTTGTCACGTTCTTGATTTGGCC	360
Db	295	GCGGCGGAAATACGGCCGCCCTTCGCAACAGGGGACAGATCTTGTCACGTTCTTGATTTGGCC	354
QY	361	GCGTACAGTGTCTTCACCCGATGGGTTGTGTGATTCGCGGACACAGTCAATTGCTTACGAG	420
Db	355	GCGTACAGTGTCTTCACCCGATGGGTTGTGTGATTCGCGGACACAGTCAATTGCTTACGAG	414
QY	421	ACCCGACAGACCGCGCGGCGGCGACACTACCCCTCTCTGGCCCGATGAGCAAGTCCCGGT	480
Db	415	ACCGTCAAGACCGCGGCGGCGGCGACACTACCCCTCTCTGGCCCGATGAGCAAGTCCCGGT	474
QY	481	CGACTCGCCGCCAGGTTGGCGCTTACCACTGTGATGCGAACCCAAAGGGGGCCGCGGTGTG	540
Db	475	CGACTCGCCGCCAGGTTGGCGCTTACCACTGTGATGCGAACCCAAAGGGGGCCGCGGTGTG	534
QY	541	CTGATGGGGCGGGGTGGCCCGCGCTGCGAACCGGGCGAGTGTGTATGCGGGCCGGCACCC	600
Db	535	CTGATGGGGCGGGGTGGCCCGCGCTGCGAACCGGGCGAGTGTGTATGCGGGCCGGCACCC	594
QY	601	GCGGCGTACAAACGCGACGCCCATCTGCGCAACAGGCGATGGGCGGACACGTTACGGTTTAAAC	660
Db	595	GCGGCGTACAAACGCGACGCCCATCTGCGCAACAGGCGATGGGCGGACACGTTAGGTTTAAAC	654
QY	661	ATCAACATCGACAAACTTGGGCACTGACGCCGGAATTCGGGCGCGGATCCACACTGCG	720
Db	655	ATCAACATCGACAAACTTGGGCACTGACGCCGGAATTCGGGCGCGGATCCACACTGCG	714
QY	721	TACTCATCGGCTCGAGCTCGAAGGGGCGCTCAAAAGTGCAGACTGGTATTTGGGGCC	780
Db	715	TACTCATCGGCTCGAGCTCGAAGGGGCGCTCAAAAGTGCAGACTGGTATTTGGGGCC	774
QY	781	GTCCTGTGCCAGGCGGCAAGGACCCCAATTAGTCTCAATTCACTTGTGCGGCATATG	840
Db	775	GTCCTGTGTGCCAGGCGGCAAGGACCCCAATTAGTCTCAATTCACTTGTGCGGCATATG	834
QY	841	AAACAGAGTGGGTAATCGGTGTGATATATAGCATGACACAGGGGGGCTGTTTGAABGGCTCA	900
Db	835	AAACAGAGTGGGTAATCGGTGTGATATATAGCATGACACAGGGGGGCTGTTTGAABGGCTCA	894
QY	901	CGACCGACCACTTACGACACCAACCGAGTTTCGCGCTGACAGACACGCTGTTTACTGCTGTG	960
Db	895	CGACCGACCACTTACGACACCAACCGAGTTTCGCGCTGACAGACACGCTGTTTACTGCTGTG	954
QY	961	GCGAATGCCCCGCTCTGCTGTGCCGAAGAGGTGACCTACGCGCTGACCAACGACGACGATG	1020

Db	995	GCNAACATGCCCGCGCTCGGTGCGCGAAGACGCTGACCTACGGCGCTGACCAACCGACGATG	1014
Qy	1021	CCGATATGTCGTCAGACTTGCCGACCATGCGTGCGGGCGCGCTGCCCGGTCAATCCGCA	1080
Db	1015	CCGATATGTCGTCGAGCTTGCCGACCATGCGTGCGGGCGCGCTGCCCGGTCAATCCGCA	1074
Qy	1081	CTACCCAAAGTCTTTCGACGCACGAAGGGCGTTACTCTCGAAGCGGTGGCCACCGAC	1140
Db	1075	CTACCCAAAGTCTTTCGACGCACGAAGGGCGTTACTCTCGAAGCGGTGGCCACCGAC	1134
Qy	1141	CTGGGGGTGCGGTTCACCGAGCCCGCAGCGTCTGCGCTGACTCTCGGCGCTCTGTATC	1200
Db	1135	CTGGGGGTGCGGTTCACCGAGCCCGCAGCGTCTGCGCTGACTCTCGGCGCTCTGTATC	1194
Qy	1201	GCCGACACACGTCGGGAGTGAAGGAAAGCATGATCTCGGC	1241
Db	1195	GCCGACACACGTCGGGAGTGAAGGAAAGCATGATCTCGNC	1235

RESULT	13
A87608	
LOCUS	A87608 1228 bp DNA PAT 22-JAN-2000
DEFINITION	Sequence 5 from Patent WO9836089.
ACCESSION	A87608
VERSION	A87608.1 GI:6736248
KEYWORDS	
SOURCE	. unidentified. unclassified.
ORGANISM	
REFERENCE	1 (bases 1 to 1228)
AUTHORS	Flohe L. and Singh M.
TITLE	TES' KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL	Pat. it: WO 9836089-A 20-AUG-1998; FLOI E LEOPOLD (DE); SINGH MAHAVIR (DE) location/Qualifiers
FEATURES	source 1..1228 /organism="unidentified" /db_xref="taxon:32644"
BASE COUNT	236 a 391 c 382 g 219 t
ORIGIN	
Query Match	97.2%; Score 1210.4; DB 5; Length 1228;
Best Local Similarity	99.4%; Pr:d. No. 1,3e-166;
Matches 1227:	Conservative 0; Mismatches 1; Indels 6; Gaps 1;
OY	1 ATCTGCAGATTAAATCGAACCTTCTTCACACTGGAAGCCTACAGTATCGAGAGGGTATC 60
Dd	1 ATTCTTGCAATTAAATCGAACCCTTTCTCATCTAGAAGCTACGATACGAGAGGGTAATTC 60
OY	61 ATGCCGTGTGATATCCACCGAGACCAAATAACAAGATTCCAATTCCGGGTGGCCATC 120
Dd	61 ATGCCGTGTGATATCCGACCCGAGACCAAAAACACG-----AATCCGGGTGGCCATC 114
OY	121 ACCCGGGCGCGGTGGCGGAACCTAACCCGTGTGGCATGAGTGTCTCATCCAGGCAGGT 180
Dd	115 ACCCGGGCGCGGTGGCGGAACCTAACCCGTGTGGCATGAGTGTCTCATCCAGGCAGGT 174
OY	181 GCCGGAAGAGGCTCGCGTATCACGACGGCGATTCCAAGGGCGGACGGCGCACTGTCTC 240
Dd	175 GCCGGAAGAGGCTCGCGTATCACCCAGCCGGATTCCAAGGGCGGCGGCGCACTGTCTC 234
OY	241 GGCACGGCGGACAGCTGTGGGCGGACCGCTGATTTATGTCTAAGGTCAAGAACCGATA 300
Dd	235 GGCACGGCGGACAGCTGTGGGCGGACCGCTGATTTATGTCTAAGGTCAAGAACCGATA 294
OY	301 GCGGCGGAATACGGCGGCTCGGACACGGCGAGATCTTGTCAGTTCCTTGCAATTGGCC 360
Dd	295 GCGGCGGAATACGGCGGCTCGGACACGGCGAGATCTTGTCAGTTCCTTGCAATTGGCC 354
OY	361 GCGTAGCTGTCTGCACCGAATGCGTTGTGGATTCGGGACACCACTCAATTGCTACGAG 420
Dd	355 GCGTAGCTGTCTGCACCGAATGCGTTGTGGATTCGGGACACCACTCAATTGCTACGAG 414

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 15:04:40 ; Search time 458.59 Seconds

(without alignments)
679.232 Million cell updates/sec

Title: US-09-362-485-2

Sequence: 1 ATCTTCGACATTAATCGAAC.....AAGCATGATGCGCCGCG 1245

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result bel., printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1245	100.0	1245	1 V49626	Mycobacterium tube
2	1245	100.0	1260	1 V49510	Mycobacterium sp.
3	1245	100.0	1260	1 V49625	Mycobacterium tube
4	467.6	37.6	682	1 V49511	Mycobacterium mari
5	261.2	21.0	1125	1 N91423	Sequence of heat-r
6	198.4	15.9	28171	1 V52155	Streptococcus pneu
7	178.4	14.3	9280	1 V74442	Streptococcus aur
8	94.8	7.6	1074	1 T67971	H. pylori membrane
9	69.2	5.6	544	1 V43039	Streptococcus pneu
10	57.4	4.6	31096	1 V74370	Staphylococcus aur
11	50.8	4.1	390	1 021833	Randomising oligon
12	50.8	4.1	390	1 036859	PCR primer for 5'
13	50.4	4.0	15872	1 T68715	Streptomyces venez
14	49	3.9	1269	1 003065	Sequence encoding
15	47.6	3.8	1413	1 X34205	Mycobacterium spec
16	47.6	3.8	1722	1 X34206	Mycobacterium spec
17	47.6	3.8	3946	1 T93610	Mycobacterium tube
18	47.6	3.8	28598	1 T89955	Sorangium cellulos
19	47.6	3.8	28958	1 T89955	Sorangium cellulos
20	47.6	3.8	49377	1 V05287	The soraphen biosy
21	46.4	3.7	985	1 V44439	Mycobacterium tube
22	46.4	3.7	985	1 V64548	M. tuberculosis im
23	46	3.7	1520	1 Q22482	grel-1 gene codin
24	46	3.7	1520	1 Q22482	grel-1 gene codin
25	46	3.7	2668	1 Q22485	grel-1 gene. Reco
26	46	3.7	17955	1 V56642	Actinoplanes sp. a
27	45.4	3.6	833	1 064203	snab gene encoding
28	45.4	3.6	5392	1 064201	Sequence comprisin
29	42.4	3.4	1833	1 064206	snR gene encoding
30	42.4	3.4	2185	1 V84066	Clone p5-4 encodin
31	42.4	3.4	2185	1 099365	S. lividans protea
32	42.4	3.4	24379	1 T93095	Streptomyces fireo
33	42.4	3.4	24379	1 V25925	Streptomyces roseo
34	42.2	3.4	329	1 V44425	Mycobacterium tube

35	42.2	3.4	329	1 V64534	M. tuberculosis im
36	42	3.4	3946	1 T93610	Mycobacterium tube
37	42	3.4	4018	1 063879	Polydactyloalkano
38	41.8	3.4	882	1 V44403	Mycobacterium tube
39	41.8	3.4	882	1 V64512	M. tuberculosis im
40	41.6	3.3	12036	1 004658	FHA structural gen
41	41.2	3.3	2414	1 005926	Sequence encoding
42	41	3.3	29879	1 046806	eryA region of S.
43	40.8	3.3	2151	1 076252	HSV-2 protease, IC
44	40.8	3.3	2151	1 076261	HSV-2 protease/ICP
45	40.8	3.3	2472	1 084671	HSV-2 UL26 gene. N

ALIGNMENTS

RESULT 1	
V49626	standard; DNA; 1245 BP.
AC V49626:	
DE 20-NOV-1998 (first entry)	
DE Mycobacterium tuberculosis L-Alanine dehydrogenase 40 KD antigen.	
KW ss: Alanine dehydrogenase; tuberculosis; TB; mycobacteria; L-alanine.	
OS Mycobacterium tuberculosis.	
PN W09836089-A2.	
PD 20-AUG-1998.	
PF 29-JAN-1998; E00483.	
PR 29-JAN-1997; EP-101338.	
PA (FLOH/) FLOHE L.	
PI Flohe L, Hutter B, Kolx A, Singh M;	
DR WPI: 98-457123/39.	
PT Diagnosis of, e.g. tuberculosis from alanine dehydrogenase activity	
PT - useful for, e.g. for diagnosis, differentiation of strains,	
PT monitoring vaccination and identification of mycobacterial	
PT inhibitors	
PS Disclosure: Fig 3.19; 55pp; German.	
CC The Mycobacterium tuberculosis L-alanine dehydrogenase (AlaDH) is used in	
CC the production of kits for diagnosing tuberculosis (TB) and other	
CC mycobacterial infections in humans or animals. Kits are used for direct	
CC diagnosis of TB on clinical samples (e.g. body fluids) and can	
CC differentiate between pathogenic and non-virulent strains, e.g. for	
CC identifying pathogens of the M.tuberculosis (M. t.) complex. The kit may	
CC also be used to identify substances that inhibit mycobacteria, for	
CC combatting epidemics and for vaccination follow-up. Oligonucleotides	
CC derived from AlaDH are used similarly in diagnostic hybridisation tests,	
CC also for culture confirmation of isolated strains and for chromosome	
CC fingerprinting to detect/differentiate between mycobacteria, and for	
CC L-alanine-specific biotransformation reactions. AlaDH is an early	
CC antigen, present extracellularly after only a few days of growth, making	
CC it an ideal drug target.	
SO Sequence 1245 BP; 238 A; 398 C; 387 G; 222 T;	

Query Match 100.0%; Score 1245; DB 1; Length 1245;
Best Local Similarity 100.0%; Pred. No. 2.3e-251;
Matches 1245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1	ATCTTCGACATTAATCGAATCTTCTTCACTGAAGCGTACGATGAGAGGGGTATC 60
DB 1	ATCTTCGACATTAATCGAATCTTCTTCACTGAAGCGTACGATGAGAGGGGTATC 60
QY 61	ATGGCGGCGGTATTCGACGAGCAACCAAAACGAATTCGAATTCGGGTGGCATC 120
DB 61	ATGGCGGCGGTATTCGACGAGCAACCAAAACGAATTCGAATTCGGGTGGCATC 120
QY 121	ACCCGCGCGGCGGTGCGGACCTAACCCGCTGCGGCAATGAGTGTCTATCGAGCGGT 180
DB 121	ACCCGCGCGGCGGTGCGGACCTAACCCGCTGCGGCAATGAGTGTCTATCGAGCGGT 180
QY 181	GCCGAGAGGGCTCGGCTATCCAGCAGCGGATTTCAAGGGGCGGCGGCAACTGTC 240
DB 181	GCCGAGAGGGCTCGGCTATCCAGCAGCGGATTTCAAGGGGCGGCGGCAACTGTC 240

OY	241	GGCACCCGCGACCAAGGATGGGGGCGGACGGCTGATTTATATGGCA-GGGCTAAAGAAACGATA	300
Db	241	GGCACCCGCGACCAAGGATGGGGGCGGACGGCTGATTTATATGGCTAAAGGTCAAGAAACGATA	300
OY	301	GGCGGCAATATACGGCGCGCTCTGGACACAGGGGACGATCTTGTTCAGGTTCTTGCAATTGGCC	360
Db	301	GGCGGCAATATACGGCGCGCTCTGGACACAGGGGACGATCTTGTTCAGGTTCTTGCAATTGGCC	360
OY	361	GGGTTCACGTGCTTGACCCATGGGTTGTGTGATTTCCGGCACACAGTCAATGGCTACAG	420
Db	361	GGGTTCACGTGCTTGACCCATGGGTTGTGTGATTTCCGGCACACGCTCAATGGCTACAG	420
OY	421	ACCGTCAGAACCGCGCGACGGCGACTACCCCTGCTTGGCCGCGATGAGCGAATGGCCGT	480
Db	421	ACCGTCAGAACCGCGCGACGGCGACTACCCCTGCTTGGCCGCGATGAGCGAATGGCCGT	480
OY	481	CGACTCGCGCGCCACAGTTGGCGCTTACCACTGATATGCCAACCCAAAGGGGGCCGGGGTGTG	540
Db	481	CGACTCGCGCGCGCCACAGTTGGCGCTTACCACTGATATGCCAACCCAAAGGGGGCCGGGGTGTG	540
OY	541	CTGATGGGCGGGGGGCGACGGCGTTCGAACCGCGCGACGTTGTGTATGAGGGCGCGGAC	600
Db	541	CTGATGGGCGGGGGGCGACGGCGTTCGAACCGCGCGACGTTGTGTATGAGGGCGCGGAC	600
OY	601	GCCGGCTACAAACGACGCCGCAATCGCCAAAGGCAATGGCGGACCGGTACGGTCTTAGAC	660
Db	601	GCCGGCTACAAACGACGCCGCAATCGCCAAAGGCAATGGCGGACCGGTACGGTCTTAGAC	660
OY	661	ATCAACATCGCAAACTTGGCAACTCGAAGCGCGAATTTCTGGGCGCGGATCCACACTCGC	720
Db	661	ATCAACATCGCAAACTTGGCAACTCGAAGCGCGAATTTCTGGGCGCGGATCCACACTCGC	720
OY	721	TACTCATCGGCTACGAGCTCGAGGGTGCCTCAAAAGTCCGACCTGTGATTTGGGGC	780
Db	721	TACTCATCGGCTACGAGCTCGAGGGTGCCTCAAAAGTCCGACCTGTGATTTGGGGC	780
OY	781	GTCCTGGTGGCCAGGCGCCAAAGGCAACCCAAATTGATCTGCATTTACTTGGCGCATATG	840
Db	781	GTCCTGGTGGCCAGGCGCCAAAGGCAACCCAAATTGATCTGCATTTACTTGGCGCATATG	840
OY	841	AAACCAAGTGGGATCTGTGTGATATAGCCATCGACACAGGGCGGCTGTTTCGAAGGCTCA	900
Db	841	AAACCAAGTGGGATCTGTGTGATATAGCCATCGACACAGGGCGGCTGTTTCGAAGGCTCA	900
OY	901	CGACCGACCACTCTAGACACGACGCTTGGCGTGCACAGACGCTGTTTACTGGGTG	960
Db	901	CGACCGACCACTCTAGACACGACGCTTGGCGTGCACAGACGCTGTTTACTGGGTG	960
OY	961	GGCAACATGCCCCGCTCGGTGCCCAAGACGTCACCTACGCGCTGACCAACGCGAGATG	1020
Db	961	GGCAACATGCCCCGCTCGGTGCCCAAGACGTCACCTACGCGCTGACCAACGCGAGATG	1020
OY	1021	CCGTATGTGCTCGAGCTTGGCGACATGGCTGGCGGGCGGCGTGGCGGTTCGAATCCGGCA	1080
Db	1021	CCGTATGTGCTCGAGCTTGGCGACATGGCTGGCGGGCGGCGTGGCGGTTCGAATCCGGCA	1080
OY	1081	CTAGCCAAAGGCTTTCTGACGCAAGAAAGGGGCTTACTGTGCGAAGGGGTGGCCACCGAC	1140
Db	1081	CTAGCCAAAGGCTTTCTGACGCAAGAAAGGGGCTTACTGTGCGAAGGGGTGGCCACCGAC	1140
OY	1141	CTGGGGGTGCGGCTTACCGAGCCCGCAGCGGTGCTGGCTGACTCTGGCGCGCTGTGTTAC	1200
Db	1141	CTGGGGGTGCGGCTTACCGAGCCCGCAGCGGTGCTGGCTGACTCTGGCGCGCTGTGTTAC	1200
OY	1201	GGCGAGCAACAGCTCGGGAGTAAAGGAAGCGATGATGTCGGCGCGC	1245
Db	1201	GGCGAGCAACAGCTCGGGAGTAAAGGAAGCGATGATGTCGGCGCGC	1245
RESULT	2		
ID	V49510	standard: DNA: 1260 BP.	
AC	V49510:		

20-OCR-1998 (first entry)
DE Mycobacterium sp. AlAdH DNA.
KM Alanine dehydrogenase: AlAdH; ADH; diagnosis; tuberculosis; pathogen;
KM swimmers' disease; vaccine; epidemic; infection; identification; ss.
OS Mycobacterium sp.
PN W09832862-A2.
PD 30-JUL-1998.
PF 29-JAN-1998; E00484.
PR 29-JAN-1997; EP-101339.
PA (FLOH/) FLOHE L.
PI Flohe L, Hutter B, Kolik A, Singh M;
PI WPI: 98-427958/36.
PT Nucleic acid encoding alanine dehydrogenase of Mycobacterium marinum
PT - used for diagnosis of tuberculosis and other mycobacterial
PT diseases; also for treatment and prevention; for drug screening and
PT for bio-transformation
PS Disclosure; Page 11; 57pp; German.
CC This sequence encodes an alanine dehydrogenase (ADH) protein isolated
CC from a Mycobacterium sp. This protein is used to diagnose tuberculosis
CC and other mycobacterial infections (including 'swimmers' disease', caused
CC by M. marinum, a fish pathogen) in humans or animals. The protein can
CC also be used for control of epidemics and for vaccination, to screen for
CC agents with anti-mycobacterial activity, and in bio-transformations that
CC are specific for L-alanine. Also mycobacteria can be identified by
CC analysis of genomic ADH sequences. ADH is an antigen that is secreted
CC early during infection.
SQ Sequence 1260 BP; 243 A; 403 C; 389 G; 225 T;

Query Match	Similarity	100.0%	Score 1245	DB 1	Length 1260
Best Local Similarity: 100.0%; Pred. No. 2, 3e-251.					
Matches 1245	Conservative	0	Mismatches	0	Indels
					Gaps
QY	1	ATCTTCGAGATTAAATCGAACTTTCTTACACTGAGCGTACGATATCGAGAGGGGTAATC	60		
DB	16	ATCTTCGAGATTAAATCGAACTTTCTTACACTGAGCGTACGATATCGAGAGGGGTAATC	75		
QY	61	ATGCGCGTCGGGATTCCGACCGAGACCAAAAACAACGAATTCGATTCGGGGTGGCATC	120		
DB	76	ATGCGCGTCGGGATTCCGACCGAGACCAAAAACAACGAATTCGATTCGGGGTGGCATC	135		
QY	121	ATCCCGCGCGCGGCTGCGGGAATTAACCCGTCGTGGCCATGAGTGTCTATCCAGGCAAGT	180		
DB	136	ATCCCGCGCGCGGCTGCGGGAATTAACCCGTCGTGGCCATGAGTGTCTATCCAGGCAAGT	195		
QY	181	GGCGGAGAGGGGCTCGGCTATACCGACGCGGATTTCAAAGGCGCGAGCGCGCAACTGGTC	240		
DB	196	GGCGGAGAGGGGCTCGGCTATACCGACGCGGATTTCAAAGGCGCGAGCGCGCAACTGGTC	255		
QY	241	GGCACCAGCGGACGAGTGTGCGGCGGATGCTATTTATTTGCTCAAGGTCAAGAAGCGATA	300		
DB	256	GGCACCAGCGGACGAGTGTGCGGCGGATGCTATTTATTTGCTCAAGGTCAAGAAGCGATA	315		
QY	301	GGGCGGGAATACGCGCGGCTGCGAGACGCGGAGATCTTGTCACAGTTCCTTGACATTGGCC	360		
DB	316	GGGCGGGAATACGCGCGGCTGCGAGACGCGGAGATCTTGTCACAGTTCCTTGACATTGGCC	375		
QY	361	GGGTCACGTGCTTGACACCGATGCGTTGTTGATTCGCGACACCACTCAATTGCTTACGAG	420		
DB	376	GGGTCACGTGCTTGACACCGATGCGTTGTTGATTCGCGACACCACTCAATTGCTTACGAG	435		
QY	421	ACCGTCGAGACCGCGAGCGGCGACATACCCCTGCTTGCGCCGATGAGCGAAGTGGCCGGT	480		
DB	436	ACCGTCGAGACCGCGAGCGGCGACATACCCCTGCTTGCGCCGATGAGCGAAGTGGCCGGT	495		
QY	481	CGAGTCGCGCGCCACAGTTGCGGCTTACCACTGATGCGAACCACAAGGGGGCGCGGTGTG	540		
DB	496	CGAGTCGCGCGCCACAGTTGCGGCTTACCACTGATGCGAACCACAAGGGGGCGCGGTGTG	555		
QY	541	CTGATGAGGCGGGGTGCGGCGCTGCAACCGGCGGCACTGTGTGATGAGGCGCGCGAC	600		
DB	556	CTGATGAGGCGGGGTGCGGCGCTGCAACCGGCGGCACTGTGTGATGAGGCGCGCGAC	615		

QY 601 GCCGCTTAACGACGACCCCGCATCCGCAACGGATGGGGCCGACCTTACGGTTCTAGAC 660
 DB 616 GCGGCTTACAGCAGCACCAGCATCCGCAACGGATGGGGCCGACCTTACGGTTCTAGAC 675
 QY 661 ATCAACATGACAAATTTGGGCAACTGAGAGCGGATGTCGGGGCGGATATCCACATCCG 720
 DB 676 ATCAACATGACAAATTTGGGCAACTGAGAGCGGATGTCGGGGCGGATATCCACATCCG 735
 QY 721 TACTATCGGCTTACGAGCTCGAGGGTGGCCGTAACAGTGGCCGACCTGTGATGGGGCC 780
 DB 736 TACTATCGGCTTACGAGCTCGAGGGTGGCCGTAACAGTGGCCGACCTGTGATGGGGCC 795
 QY 781 GTCTGTGTCGAGGCGGCAAGGCAATTTAGTCTCGAATTCACCTTGTGGGCGCATATG 840
 DB 796 GTCTGTGTCGAGGCGGCAAGGCAATTTAGTCTCGAATTCACCTTGTGGGCGCATATG 855
 QY 841 AAACGAGTGGCTTACGAGCTCGAGGGTGGCCGTAACAGTGGCCGACCTGTGATGGGGCC 900
 DB 856 AAACGAGTGGCTTACGAGCTCGAGGGTGGCCGTAACAGTGGCCGACCTGTGATGGGGCC 915
 QY 901 CGACGACGACCTTACGAGCTCGAGGGTGGCCGTAACAGTGGCCGACCTGTGATGGGGCC 960
 DB 916 CGACGACGACCTTACGAGCTCGAGGGTGGCCGTAACAGTGGCCGACCTGTGATGGGGCC 975
 QY 961 GCGAATGTCGAGGCTGTCGAGGGTGGCCGTAACAGTGGCCGACCTGTGATGGGGCC 1020
 DB 976 GCGAATGTCGAGGCTGTCGAGGGTGGCCGTAACAGTGGCCGACCTGTGATGGGGCC 1035
 QY 1021 CGCTATGTCGAGGCTGTCGAGGGTGGCCGTAACAGTGGCCGACCTGTGATGGGGCC 1080
 DB 1036 CGCTATGTCGAGGCTGTCGAGGGTGGCCGTAACAGTGGCCGACCTGTGATGGGGCC 1095
 QY 1081 CTAGCCAAAGTCTTTCGAGCAGCAAGAGGGCGTTACTGTCCGAAGCGGTGGCCACCGAC 1140
 DB 1096 CTAGCCAAAGTCTTTCGAGCAGCAAGAGGGCGTTACTGTCCGAAGCGGTGGCCACCGAC 1155
 QY 1141 CTGGGGGTGCTTTCGAGCAGCAAGAGGGCGTTACTGTCCGAAGCGGTGGCCACCGAC 1200
 DB 1156 CTGGGGGTGCTTTCGAGCAGCAAGAGGGCGTTACTGTCCGAAGCGGTGGCCACCGAC 1215
 QY 1201 GCGGAGCAGCTGGGGAGTAAGGAAGCATGTGTCGGCCGG 1245
 DB 1216 GCGGAGCAGCTGGGGAGTAAGGAAGCATGTGTCGGCCGG 1260
 RESULT 3
 V49625
 ID V49625 standard: DNA: 1260 BP.
 AC V49625;
 DT 20-NOV-1998 (first entry)
 DE Mycobacterium tuberculosis: Alanine dehydrogenase.
 KW ss: Alanine dehydrogenase; tuberculosis; TB; mycobacteria; L-alanine.
 OS Mycobacterium tuberculosis.
 PN M09836088-A2.
 PD 20-AUG-1998.
 PR 29-JAN-1998; E00483.
 PA (FLOH) FLOHE L.
 PI Flohe L, Hutter B, Kolk A, Singh M;
 DR WPI: 98-457123/39.
 PT Diagnosis of, e.g. tuberculosis from alanine dehydrogenase activity
 PT useful for, e.g. for diagnosis, differentiation of strains,
 PT monitoring vaccination and identification of mycobacterial
 PT inhibitors
 CS Claim 13: Fig 2.3; 55pp; German.
 CC The Mycobacterium tuberculosis alanine dehydrogenase (AlaDH) is in
 CC the production of kits for diagnosing tuberculosis (TB) and other
 CC mycobacterial infections in humans or animals. Kits are used for direct
 CC diagnosis of TB on clinical samples (e.g. body fluids) and can
 CC differentiate between pathogenic and non-virulent strains, e.g. for
 CC identifying pathogens of the M. tuberculosis (M. t.) complex. The kit may
 CC also be used to identify substances that inhibit mycobacteria, for
 CC combatting epidemics and for vaccination follow-up. Oligonucleotides

CC derived from AlaDH are used similarly in diagnostic hybridisation tests,
 CC also for culture confirmation of isolated strains and for chromosome
 CC fingerprinting to detect/differentiate between mycobacteria, and for
 CC L-alanine-specific biotransformation reactions. AlaDH is an early
 CC antigen, present extracellularly after only a few days of growth, making
 CC it an ideal drug target.
 SQ Sequence 1260 BP; 243 A; 403 C; 389 G; 225 T;
 Query Match 100.0%; Sec: 12.5; DB 1; Length 1260;
 Best Local Similarity 100.0%; Pred. No. 2.3e-251;
 Matches 1245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATCTTGCAATTAATGCACTTTCTTACACTGAGACCGTACAGTATCGAGAGGGTATTC 60
 DB 16 ATCTTGCAATTAATGCACTTTCTTACACTGAGAGCGTACAGTATCGAGAGGGTATTC 75
 QY 61 ATGCGGCTGGTATTCGACCGAGACCAAAACAGATTCATTCGGGGTGGCCATC 120
 DB 76 ATGCGGCTGGTATTCGACCGAGACCAAAACAGATTCATTCGGGGTGGCCATC 135
 QY 121 ACCCGGCGGCGTGGGAGTAAACCGTGTGGCCATGAGGTGCTATCCAGCAGGT 180
 DB 136 ACCCGGCGGCGTGGGAGTAAACCGTGTGGCCATGAGGTGCTATCCAGCAGGT 195
 QY 181 GCGGAGAGGCTGCGCTATACCGAGCGGATTTCAAGCGGCGGAGCGGCACTGTC 240
 DB 196 GCGGAGAGGCTGCGCTATACCGAGCGGATTTCAAGCGGCGGAGCGGCACTGTC 255
 QY 241 GGCACCGCGGACAGGTGTGGGGCGAGCGTGTATTTCTCAAGGTCAAAAGCGATA 300
 DB 256 GGCACCGCGGACAGGTGTGGGGCGAGCGTGTATTTCTCAAGGTCAAAAGCGATA 315
 QY 301 GCGGCGGAATACGCGCGCTGCGACAGGGGAGATTTGTACGTTCTTGCATTTGGCC 360
 DB 316 GCGGCGGAATACGCGCGCTGCGACAGGGGAGATTTGTACGTTCTTGCATTTGGCC 375
 QY 361 GCGTCACTGCTTTCGACCGCATGCTGTGTGATTCGGGACCACTCATTTGCGTACGAG 420
 DB 376 GCGTCACTGCTTTCGACCGCATGCTGTGTGATTCGGGACCACTCATTTGCGTACGAG 435
 QY 421 ACCGTCCAGACCGCGAGCGC :CACTACCGCTGCTGCCCGATGAGCGAAGTGGCGGT 480
 DB 436 ACCGTCCAGACCGCGAGCGC :CACTACCGCTGCTGCCCGATGAGCGAAGTGGCGGT 495
 QY 481 CGACTGCGCGCGGCTTGGCTTACCACTGATGCGAAGCCAGAGGGGCGGCTGTG 540
 DB 496 CGACTGCGCGCGGCTTGGCTTACCACTGATGCGAAGCCAGAGGGGCGGCTGTG 555
 QY 541 CTGATGGGGGGGTGCTGCTGCGAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 DB 556 CTGATGGGGGGGTGCTGCTGCGAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 615
 QY 601 GCGGCGCTACAAAGCGAGCGGCTGCGAAGCGGATGGGGCGGCGGCTTACGAG 660
 DB 616 GCGGCGCTACAAAGCGAGCGGCTGCGAAGCGGATGGGGCGGCGGCTTACGAG 675
 QY 661 ATCAACATGACAAATTTGGGCAACTGAGAGCGGATGTCGGGGCGGATATCCACATCCG 720
 DB 676 ATCAACATGACAAATTTGGGCAACTGAGAGCGGATGTCGGGGCGGATATCCACATCCG 735
 QY 721 TACTATCGGCTTACGAGCTCGAGGGTGGCCGTAACAGTGGCCGACCTGTGATGGGGCC 780
 DB 736 TACTATCGGCTTACGAGCTCGAGGGTGGCCGTAACAGTGGCCGACCTGTGATGGGGCC 795
 QY 781 GTCTGTGTCGAGGCGGCAAGGCAATTTAGTCTCGAATTCACCTTGTGGGCGCATATG 840
 DB 796 GTCTGTGTCGAGGCGGCAAGGCAATTTAGTCTCGAATTCACCTTGTGGGCGCATATG 855
 QY 841 AAACGAGTGGCTTACGAGCTCGAGGGTGGCCGTAACAGTGGCCGACCTGTGATGGGGCC 900
 DB 856 AAACGAGTGGCTTACGAGCTCGAGGGTGGCCGTAACAGTGGCCGACCTGTGATGGGGCC 915

Matches 604; Conservative 0; Mismatches 458; Indels 18; Gaps 5;

```

QY 61 ATGCGGCTGATATTCGACCGAGACCAAAACAGATTCATTCGGGGTCCATC 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 ATGAAGATCGGATTCGACCAAAATCAAAACATG-----AAAACCGCGGCATC 54
QY 121 ACCCGGGGGGGGGGGGAAACCAACCGGCTGGGCGCATGAGGGGCT---CATCCAGCA 177
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 55 ACTCGGCGAGCGGTGATGACGCTGCTAAACCGGGGCGATAGGTGGTGTATGGAGACG 114
QY 178 GGTGCGGAGAGGGCTGCGCTATCCACGACGCGGATTTCAAGGGCGGCAAGCGCACTG 237
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 115 GAAGGGGGGGCTGGGTGGGGGTTTCCGATTCGAGATGAAAAAGCGGGGAGCTGAT 174
QY 238 GTGCGGACCGCGGACGACGAGTGGGGCGGAGCTATTTATTCGCAAGCTCAAGAACG 297
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175 CATTGCGGAACGTGGAGATGCTTGGAGCGGGGAGATGGTTGAAAGTGAAGAGCGG 234
QY 298 ATACGGCGGAATACGGCGCGCTCGACAGCGGAGATCTTGTTCAGCTTCTTGATTTG 357
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 235 CTGGCTGAGAGATTCGCTATTTTCGCGCGGATGATTTTGTATGATTTTCATTTA 294
QY 358 GCCGCGTACGCTGCTGACCGGATCGTTGATTCGCGACCGCTCATTCCTTAC 417
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 295 GCCGCGCGGCGAAGCGCTACGAAAGCGCTGCTGAGCAAAAAGTGTGCGCATGCTTAC 354
QY 418 GAGACCTCCAGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 355 GAGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 414
QY 478 GGTGACTCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 415 GCGCGCATGTGGGTGGAGTGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 474
QY 538 GTGTGATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 597
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 475 ATTTGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 534
QY 598 ACCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 657
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 535 ACCGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 594
QY 658 GACATCAACATCGACMACTTCCGACACTGACGCGGCGGCGGCGGCGGCGGCGGCGG 717
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 595 GACATTAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 654
QY 718 CGCTACTATGCGGCTAGAGACTCGAGGCTGCCGCAACGTCGCGGCGGCGGCGGCGG 777
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 655 CTGATGTCACACTGCTACCATATTCGCGAGTGCCTGCGGCGGCGGCGGCGGCGG 714
QY 778 GCCGCTGCTGTCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 837
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 715 GCCGCTGCTGATCCGGGGGCGAA---AGGAAAGTGTGAGGAAAGATGATGGCTCG 771
QY 838 ATGAACAAGAGTGGCGGCTAGTGTGATATAGCATTCAGACAGGCGGCTGTTTCAA--- 895
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 772 ATGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 831
QY 895 GAGCTACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 954
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 892 ACCGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 891
QY 958 TGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1014
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 892 GCCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 948
QY 1015 ACGATGCGGCTATGCTCGAGCTGCGACGACGACGACGACGACGACGACGACGACGAC 1074
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 949 ACGATGCGGCTATGCTCGAGCTGCGACGACGACGACGACGACGACGACGACGACGAC 1008
QY 1075 CCGGACGCTAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1134
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1009 CCGGCGGCTGTTTAAAGGATCAACAGGCTGCGACGCGGCGGCGGCGGCGGCGGCGG 1068

```

RESULT 6
V52135/c
ID V52155 standard; DNA; 28171 BP.
AC 23-OCT-1998 (first entry)
DE Streptococcus pneumoniae genome fragment SEQ ID NO:22.
KW Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
KM computer readable medium; vaccine; pharmaceutical composition; ds.
OS Streptococcus pneumoniae.
PN WO9818931-A2.
PD 07-MAY-1998.
PE 30-OCT-1997; 019588.
PR 31-OCT-1996; US-029960.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M.
PI Kunsch CA, Rosen CA;
PI MPI: 98-272225/24.
PT Computer-readable medium with recorded Streptococcus pneumoniae
PT polynucleotide sequences - useful in diagnostic kits and assays, and
PT pharmaceutical compositions and vaccines for Streptococcus
PT pneumoniae.
PS Claim 1, Page 273-289; 1409pp; English.
CC The present invention describes a computer readable medium which has
CC the nucleotide sequences SEQ ID NO:1 to 391 (V52134 to V52524) recorded
CC on it, or a representative fragment or a sequence at least 95% identical
CC to SEQ ID NO:1 to 391. The nucleotide sequences depicted in SEQ ID NO:1
CC to 391 (V52134 to V52524) are genomic fragments from Streptococcus
CC pneumoniae. The present invention also describes an isolated nucleic acid
CC molecule encoding a homologue of any of the fragments of the S. pneumoniae
CC genome (SEQ ID NO:1 to 391) where the nucleic acid molecule is produced
CC by a process comprising: (a) screening a genomic DNA library using as a
CC probe a target sequence defined by any of the sequences in SEQ ID NO:1
CC to 391, identifying members of the library which contain sequences
CC that hybridize to the target sequence and isolating the nucleic acid
CC molecules from the members; or (b) isolating mRNA, DNA or cDNA produced
CC from an organism, amplifying nucleic acid molecules whose nucleotide
CC sequence is homologous to amplification primers derived from the
CC fragment of the S. pneumoniae genome to prime the amplification and
CC isolating the amplified sequences. The computer readable medium can be
CC used in a computer-based system for identifying fragments of the
CC S. pneumoniae genome of commercial importance, or expression modulating
CC fragments of the S. pneumoniae genome. Products from the present
CC invention can be used in diagnosis kits and assays, and pharmaceutical
CC compositions and vaccines for S. pneumoniae.
SQ Sequence 28171 BP; 8507 A; 5080 C; 6440 G; 8144 T;

Query Match 15.9%; Score 198.4; DB 1; Length 28171;
Best Local Similarly 53.4%; Pred. No. 3.7e-33;
Matches 516; Conservative 0; Mismatches 431; Indels 19; Gaps 4;

```

QY 89 AAACAACGATTCATTCGCGGCTGCGCATACCGCGCGCGGCGGCGGCGGCGGCGGCGG 148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21689 AAATTTAAATAACGAAACCGTTCGCTCACACTGACGATGCTCATAGCTAGTTA 21630
QY 149 GTCTGGCCATGAGCTGCTCATCAGGACGAGTGGCGGAGAGGGCTCGCTATACGACG 208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21629 GTCTGGTCATCGTGTCTTATCGAAACAAATGCTGTCTGCTTCTGGCTTACGATG 21570
QY 209 CGGATTCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21589 CTGACTATCAAAACAGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 21513
QY 269 CTGATTTATTCCTAAGGTCAAAACGATAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21512 CAGAGTTGGTGTGAAGTAAAGAACTTTAAGTTCTGATAGGTTACTTGGCGGACG 21453
QY 329 GCGAGATCTTGTACGCTTCTGATTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21452 ATCTTCTCTCTTCACTTCTGACATGCGGCTGCTCCAGATTTAGCAGATGCTATGT 21393

```

OY		389	TGATNTCCGGACACAGCCTTGGCTTCAGAGACCCTGCAGACCGCCGACGCCGACTAC	448
OY		21392	T-----AACAGCAAAAACAACACTGAACCTGTTCCGTGACAAATCAGAGCAAACTAC	21345
OY		449	CCCTGCTTGCCCCGATGAGCGAAGTCGCCGGTGCAGCTCCGCCGCCAGGTGGCGCTTACC	508
Db		21344	CGCTTCCTGTTCTCATAGTAGTAGAGTTGCAGAGTCGTAATGGCTGTTCAAATCCGACCTCACT	21285
OY		509	ACCTGATGCGAACCBAAGGGGGCCCGGGTGTGCTATGAGGGGGGGGTGGCCGGCGTCGAAC	568
Db		21284	TTCCTTACTAAGCAAGCTGTGGCTCTGTGTTCTTACTTGGTGGTGTACACAGSTGTTCCAA	21225
OY		569	CGCGCGAGCTGTGGTGATGTCGGCGCGCGCACCGCGCGGTACAAACGACGCCGCATCGCCA	628
Db		21224	AAGAAAAAGTAACATATCATCCGGTGGTGGTGTGCTGTGATACATCTCGTCCGCATCGCCC	21165
OY		629	ACGGCATGGGGCGGACCGTTACGCTTCTAGCATCAACATCGACAAACTTGGCAATCG	688
Db		21164	TTGGCTTGGTGCTCAAGTGAAGCTATTTTAGATATTAGTTCCAAAGCGTCTTCAGTTTTAG	21105
OY		689	ACGGCAGTTCGTGGCGCGGATCCACACTCGCTACTCATCGGCTTACGAGTCTCGAGGGTG	748
Db		21104	AAGAAGTCTTTGGAGTGAAGTCAAAATCAACICTTATGTCTAATTCATTCAACATTAAGCAA	21045
OY		749	CCGTCAAACGTGCCGACCTGTGATTTGGGGCGCTCCTGCTC CAGCGCCAAAGCACCA	78
Db		21044	GTTGGAGATGCTCTATGTGGGATTTGGAGCCATTCTCATCCCTGGTGMAAAGCACCGG	20985
OY		809	AATTGTCGCAATTCACCTGTGCGGCAATCAAAACAGGTCCGGTACGTGGATATAG	868
Db		20984	AATTGTGACAGATAGATGATGTCAAACAATAATCGTCCAGGCTCTGTGA -TCGTTGACGTTG	20926
OY		869	CCATCGACCAAGGCGCGCTGTTTGA --AGGCTCAGACGACCAACCTAAGCACCCACGA	925
Db		20925	CTGTGACCAAGTGGCGGTTATCGAAACAGCTGACCCGTGTCAACAGCACAGTAGAACCGG	20866
OY		926	CGTTGCGCGTGCAGACAGACGCTGTTTACTGGGTGGCGAACAATGCCCGCTCGGTGCGCA	985
Db		20865	TCTATGA AAAAACAGCGTGTCTTCACATAGCCGTTGCCAATATCCCTGGTGGGTGCTC	20806
OY		986	AGACGCTGCACCTACCGCTGCAGCAACGCGAGCATGCCGTATGCTGCAGCTTGGCCGACC	1045
Db		20805	GCACCTCAACCATCGCGCCTAACACATGTGACATCTTCTATTATCGAAGCTTTGGTGCGA	20746
OY		1046	ATGGCT 1051	
Db		20745	AAGGAT 20740	
RESULT		7		
V74442/c				
ID		V74442	standard; DNA; 9280 BP.	
AC		V74442:		
DC		16-MAR-1999	(first entry)	
DE		Staphylococcus aureus contig SHQ ID #131.		
KM		Computer readable medium; vaccine; S.aureus infection; immunodetection;		
KW		cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;		
KX		skin infection; surgical wound infection; scalded skin syndrome;		
KZ		toxic shock syndrome; ds.		
OS		Staphylococcus aureus.		
FH		Key	Location/Qualifiers	
FT		misc_feature	841..900	
FT			/tag=a	
FT			/note=	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
FT		misc_feature	2641..2700	
FT			/tag=b	
FT			/note=	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

Query Match	Best Local Similarity	Score	178.4	DB 1	Length	9280
Matches	484	Conservative	0	Mismatches	466	Indels
157	CATGAGGTGCTATCCAGCAGCGTCCGGAGAGGCTCGCATATCCAGCCGGATTTC	216				
9280	CATGAGGTGCTATCCAGCAGCGTCCGGAGAGGCTCGCATATCCAGCCGGATTTC	9221				
217	AAGCGCGAGCGCGCACTGTCGCGACCCGCCACAGGTGTGGCGCAGCGTATTTA	276				
9220	AAGCGCGAGCGCGCACTGTCGCGACCCGCCACAGGTGTGGCGCAGCGTATTTA	9164				
277	TGAGCTCAAGCTCAAGAACGATGAGGGGGGGAATACGGCGCTCGCGACGAGCGGATC	336				
9163	GTTATTAAAGTTAAAGAACCACTTGATATGATATATTCATATTTTAAAGAAAGGCTTGTGA	9104				
337	TGTTTCACGCTTTCGATTTGGCGCGGTGACGATGCTTTCACGCGATGCTGTTGATTC	396				
9103	TGTTTCACGCTTTCGATTTGGCGCGGTGACGATGCTTTCACGCGATGCTGTTGATTC	9044				
397	GGCACCAGTCATTTGCTTACGAGACCGTCCAGACCGCGGACGCGCTACCCCTGCTT	456				


```

Db 9043 AAGTAATTAGTATGTCATATGAGACTGTGCAGTACCAAGACCATTTTACATGTTA 8984
Qy 457 GCGCCGATGACGAGCAAGTCGCCGCGTCGATCCGCCCGAGTTGGCGCTTACCACCTATG 516
Db 8983 TCACCAATGAGTGAAGTAGACGAGAAATGTACGCTCAAGTTGGCGAGAGTTCCCTACAA 8924
Qy 517 CGAACCCCAAGGGGCGCGGTGCTGATGGGCGGGGCGCGCGCGGTGAGAACGGCGGAC 576
Db 8923 AACTTATATGTTGGTATGAGTAATCTACTAGTGTGTCCAGAGAGTACCTAAGGTTAAA 8864
Qy 577 GTCTGTGATGCGCGCGCGGACCGCGGCTACAAACGACGCGCCCATCGCAACGGCATG 636
Db 8863 GTACTATATGCTGGTGGTGTGCAAGCAGAACAAATGACGCTAAATATGCACTAGACTA 8804
Qy 637 GGGCGACCGCTTACGCTGCTAGACATCAATCAACAACTTCGGCACTGAGCGCCAG 696
Db 8803 GGTCCAGATGTTACGATTTTGAATGTTATCCAAAGCGTTTACCAACAAATGATGATTTA 8744
Qy 697 TTCTGCGCGCGGATCCACACTGCTACTCATCGGCTTACGAGCTCGAGGTTGCGGTCAA 756
Db 8743 TTGCGTGACGTGATACATCAATATGTCAAATCCGTTGAAATGATGATGATGATTTAA 8684
Qy 757 CGTGCGCACTGTGATGTTGGCGCGTCTGTGTCAGGCGCCAGGACCAACCAATTAGTC 816
Db 8683 CAAAGTATTTAGTAATAGTGTGCAAGTATTTAATCCAGGTGCTAAAGCCCAAGACTTGT 8624
Qy 817 TCGAATTACTGTGCGCATATGAAACAGGTGGGCTGAGTGGTATATGACCATGAC 876
Db 8623 ACAGAACATGATTAATCAATATGAAATGAGTGTGATTAATGATGATGATGATGATTAAT 8564
Qy 877 CAGGCGCGCTGTTCGAAAGGCTC--ACGACCGACCATGACGACCGGCTGCGC 933
Db 8563 CAAAGCGGATTTTGAACAAAGATGATAAATTCAGACATGATGATGATGATGATGATTAAT 8504
Qy 934 GTGACGACGACGCTGTTTACTGCGTGGCGCAATCCCGCTGGTGGCGCAAGCTGCG 993
Db 8503 AAGCATGTGTGTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8444
Qy 994 ACCTACGCGGTGACCAAGCGACGATGCGTATGCTGAGCTGAGCTGGCGGACCATGGCTG 1053
Db 8443 ACCTTACGTTTAAATATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATG 8384
Qy 1054 CGGCGCGCTGCGGCTGCAATCCGCTACCTAGCCAAAGGCTTTCGACCGACGAGG 1109
Db 8383 AGAGAGCATTTTAATCAATCAACCATATCATAGTTTAATTAATCAATCAAGG 8328

RESULT 8
ID T67971 standard: DNA: 1074 BP.
AC T67971:
DE 15-JUN-1997 (first entry)
KW H. pylori membrane protein ORF 05cp20518orf61.
KW Vaccine; prevention; treatment; infection; identification;
KW binding compound; bacterium; life cycle; activator; bacteria;
KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis;
KW membrane; amino acid; metabolism; ds.
OS Helicobacter pylori.
FH key Location/Qualifiers
FT cds 1..1074
FT /*tag- a
FT /note- "no stop codon given"
PN MO9640893-A1.
PD 19-DEC-1996.
PE 06-JUN-1996: U09122.
PR 07-JUN-1995: US-487032.
PR 01-APR-1996: US-630405.
PA (ASTR ) ASTRA AB.
PI Berglundh O, Smith D, Mellgaerd BL;
DR P-PSDB: W20718.
PT Helicobacter pylori nucleic acid sequences and related
polyptide(s) - useful for vaccines to treat or prevent H. pylori

```

```

PT infection, and to detect Helicobacter
PS Claim 27: Page 819: 1481bp: English.
CC The present sequence encodes a Helicobacter pylori membrane
CC protein likely to contain four membrane spanning regions.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds,
CC useful as potential H. pylori life cycle activators or inhibitors.
CC The genomic sequence of H. pylori (ATCC 55679) was determined from
CC overlapping contigs generated by mechanically shearing the
CC bacterial DNA. The sequences were analysed for ORF of at least 180
CC nucleotides, and the predicted coding regions defined by computer
CC evaluation. To identify likely H. pylori antigens for vaccine
CC development, the amino acid sequences predicted from various ORF
CC were analysed for significant homology to other known or exported
CC membrane proteins. Having identified and determined the sequences
CC of interest, particular regions can be isolated from H. pylori by
CC PCR amplification for recombinant polypeptide production, e.g. in
CC E. coli hosts.
SQ Sequence 1074 BP: 302 A: 189 C: 297 G: 286 T:

```

```

Query Match 7.68: Score 94.8: DB 1: Length 1074:
Best Local Similarity 46.18: Pred. No. 1.1e-11:
Matches 370: Conservative 0: Mismatches 417: Indels 15: Gaps 1:

```

```

Qy 271 GATTATTCGTCAGTCGTAAGAACCGATAGCGCGGAAATCGGCGCTGCGACAGCG 330
Db 133 GATTGTTGGTGAATGCAAGAGCCCTTTAGACCATGATCCCTTTGCTCAAGAAAA 192
Qy 331 CAGATCTGTGTCAGTCTTCTGATTTGGCGCGTACAGTGTCTGACGATGCTGTG 390
Db 193 GCGACTGCTGTTATTTATTTAGATTTAGCTATCAAAAAAGCTTGTGTAATTTTATT 252
Qy 391 GATTCGGGACACGTCATTCCTACGAGACCGTCCAGACGCGCGGACGCTACCC 450
Db 253 AATAAAAAATCACTTCTATTTGCACTGAACCAATGCGCGGCTCAAAAGCATACCT 312
Qy 451 CTGCTGCGCGGATGAGGAGTGGCGGTGATCGCGCGCGCGCGCGCGCGCGCGCG 502
Db 313 ATTTAGCGCTTATGAGGTGTGTGGTGGAGGTGGTGGCATTTATGTCACACATTAAT 372
Qy 502 -----GCTTACCACTGATGCGAACCCCAAGGGCGCGGTGATGATGGGGGGTG 555
Db 373 TTACTGCTTTAGACATGTTAAAGGTTTATGGGTAAGGGGTCATGCTATGAGGGTTTG 432
Qy 556 CCGGCGTGAACCGCGGAGCTGTGTATCGCGCGCGCGCGCGCGCGCGCTACAGCA 615
Db 433 TCGGTTGCGCAAGGGCTAAATCGTCTATTTGAGGCGGTGTGGCATGAGAGC 492
Qy 616 GCGCGCATCGGCAAGCGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 675
Db 493 GCGAAGGCTTAAAGCCAAATGCGGGCTAAAGTAAAGTATTTAGATTAAGTTAGCTAA 552
Qy 676 CTGCGCACTGACGCGGAGTCTTGGCGCGGATCCACACTGCTACTGATGCGCTAC 735
Db 553 TTACAAAACACCTTATATCATTTTATATTTAGAAAGCTTAAAGCGTAAAGAACCC 612
Qy 736 GAGCTGAGGCTGCCGTCAAAGTGGCGACCTGTGATTTGGGCGCTGTGTCAGGC 795
Db 613 AATATCATTAAGCCTTAAAGCGGCGGTAGGAGGCGGTGCTGCTGATTCAGCG 672
Qy 796 GCCAAGGACCCCAATTTGCTCGAATCACTTGTGCGGCTATGAAACGAGTGGCGTA 855
Db 673 AGCCAAACCCCTTAAAGTATCTTAAAGAAAGATTTAAATCATGACGACCAAGGGGTA 732
Qy 856 CTGCTGATATAGCATGACGACGAGGCGCTGTTGCAAGGCTCAGCAGCAGCACCATC 915
Db 733 GTCAATGATGTGCTTGTGAGTTAGGGGGGTCATTAAGCATACGCAAGCAAGCAT 792
Qy 916 GACCAACCGAGTTCGCGGTGACGACAGCTGTTTACTGCTGCGGAGCATGCGCGCC 975
Db 793 TCTAACCGGCTGTGTGAGAAAGTTTGTTCATATATGCGTGGCAACATCCAGGG 852

```


QY 976 TCGGTGGCGAAGACGTGCACCTACGGCGGTGACCACGACCATGGCGATGTGCTCAG 1035
 Db 853 ATTGTCCCTAAAGACGACCTACGGCTTATAGCCATGCGAGTGTCCGATTTGTTGAT 912
 QY 1036 CTTCGCCACCATGGCTGGCGGG 1057
 Db 913 TATTTAGAGCATGGCTGGAAGG 934

VS	RESULT	9
AC	143039/C	
ID	V43039 standard; DNA: 544 BP.	
DT	09-NOV-1998 (first entry)	
DE	Streptococcus pneumoniae polypeptide coding region.	
KW	Polypeptide; ORF; open reading frame; infection; bacterial;	
KW	Streptococcal; Dactereamia; diagnosis; prophylaxis; ds.	
OS	Streptococcus pneumoniae.	
FH	Key	Location/Qualifiers
FT	complement (39..334)	
FT	/*tag= a	
FT	/note= "polypeptide"	
PN	M09823631-A1.	
PD	04-JUN-1998.	
PP	24-NOV-1997; U21975.	
PR	27-NOV-1996; US-031879.	
PA	(SMIK) SMITHKLINE BEECHAM CORP.	
PA	(SMIK) SMITHKLINE BEECHAM PLC.	
PI	Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO,	
PI	Reid RH, Zarfos PN;	
PI	WPI: 98-322654/28.	
DR	P-BSD3; M62769.	
PT	Streptococcus pneumoniae polynucleotides - useful for developing	
PT	products for diagnosis, prevention and treatment of infections e.g.	
PT	pneumonia, bacteraemia, meningitis or endocarditis	
PS	Claim 1; Page 165-166; 181pp; English.	
CC	The sequence is that of a Streptococcal polypeptide coding region.	
CC	The polypeptide can potentially be used for the diagnosis and	
CC	prevention of bacterial infections, especially SP infection.	
CC	It may be used for the treatment of diseases such as otitis media,	
CC	conjunctivitis, pneumonia, bacteraemia, meningitis, sinusitis, pleura	
CC	empyema, endocarditis or infection of the cerebrospinal fluid.	
SO	Sequence 544 BP; 149 A; 121 C; 124 G; 150 T;	

RESULT	10
ID	V74370
AC	V74370 standard; DNA; 31096 BP.
DT	16-MAR-1999 (first entry)
DE	Staphylococcus aureus contig S8Q ID #59.
KM	Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW	cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW	skin infection; surgical wound infection; scalded skin syndrome;
KW	toxic shock syndrome; ds.
OS	Staphylococcus aureus
RH	Key
RH	Location/Qualifiers
FT	misc_feature
FT	1201..1260
FT	/tag-a
FT	/note-
FT	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
FT	3001..3060
FT	/tag-d
FT	/note-
FT	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
FT	4801..4860
FT	/tag-c
FT	/note-
FT	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
FT	6601..6660
FT	/tag-d
FT	/note-
FT	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
FT	8401..8460
FT	/tag-e
FT	/note-
FT	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
FT	10201..10260
FT	/tag-f
FT	/note-
FT	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
FT	12001..12060
FT	/tag-u
FT	/note-
FT	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
FT	13801..13860
FT	/tag-h
FT	/note-
FT	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
FT	15601..15660
FT	/tag-i
FT	/note-
FT	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
FT	17401..17460
FT	/tag-j
FT	/note-
FT	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
FT	misc_feature
FT	17401..17460
FT	/tag-k
FT	/note-
FT	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 14:58:42 ; Search time 341.15 Seconds

(without alignments)
474.370 Million cell updates/sec

Title: US-09-362-485-2

Perfect score: 1245
Sequence: 1 ATCTGCAGATATATCGAAC.....AACGATGATGCGCCGCG 1245

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/5C.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/5D.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/6.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/PTCUTS.COMB.seq:*
7: /cgn2_6/ptodata/2/ina/Dackfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47.6	3.8	1620	3	US-08-461-775-10
2	47.6	3.8	28958	1	US-08-258-261B-6
3	47.6	3.8	28958	1	US-08-456-837-6
4	47.6	3.8	28958	1	US-08-457-342-6
5	47.6	3.8	28958	1	US-08-457-646A-6
6	47.6	3.8	28958	2	US-08-458-076A-6
7	47.6	3.8	28958	2	US-08-764-233A-4
8	47.6	3.8	28958	2	US-08-457-335A-6
9	47.6	3.8	28958	2	US-08-729-214-6
10	47.6	3.8	49377	2	US-08-764-233A-1
11	46	3.7	2582	4	US-08-816-105A-2
12	45.4	3.6	833	3	US-08-403-852D-3
13	45.4	3.6	5392	3	US-08-403-852D-1
14	42.4	3.4	1833	3	US-08-403-852D-6
15	42.4	3.4	2185	1	US-08-173-508-3
16	42.4	3.4	2185	3	US-08-265-310-3
17	41.2	3.3	2414	7	US-08-461-775-11
18	41.2	3.3	2668	3	US-07-642-734C-3
19	41.2	3.3	20235	2	US-08-804-227C-7
20	40.8	3.3	44377	3	US-08-804-158-1
21	40.8	3.3	44377	4	US-08-690-473-1
22	40.6	3.3	13001	2	US-08-458-568A-11
23	40.2	3.2	2064	1	US-08-343-428-1
24	40.2	3.2	11219	2	US-07-642-734C-1
25	39.6	3.2	474	3	US-08-403-852D-14
26	39.4	3.2	459	4	US-08-387-942C-35
27					

C 28	39.4	3.2	1998	7	5212296-8	Patent No. 5212296
C 29	39.4	3.2	12588	4	US-08-387-942C-1	Sequence 1, Appl
C 30	39.2	3.1	420	1	US-08-470-179-148	Sequence 148, App
C 31	39.2	3.1	2353	6	PCT-US92-05840-1	Sequence 1, Appl
C 32	39.2	3.1	43280	3	US-08-804-227C-1	Sequence 1, Appl
C 33	38.8	3.1	1524	7	5512669-1	Patent No. 5512669
C 34	38.8	3.1	1525	7	5229279-1	Patent No. 5229279
C 35	38.8	3.1	8051	4	US-08-576-626A-2	Sequence 2, Appl
C 36	38.8	3.1	11219	2	US-07-642-734C-1	Sequence 2, Appl
C 37	38.6	3.1	1187	2	US-08-440-856A-2	Sequence 2, Appl
C 38	38.6	3.1	1231	2	US-08-074-121-4	Sequence 4, Appl
C 39	38.6	3.1	3231	6	PCT-US94-06447-4	Sequence 4, Appl
C 40	38.4	3.1	1215	4	US-08-947-726A-1	Sequence 20, Appl
C 41	38.2	3.1	2109	4	US-08-555-568B-20	Sequence 1, Appl
C 42	38	3.1	30001	1	US-08-125-458-1	Sequence 1, Appl
C 43	38	3.1	30001	3	US-08-474-933-1	Sequence 1, Appl
C 44	37.8	3.0	2588	3	US-08-796-414B-6	Sequence 6, Appl
C 45	37.8	3.0	15664	1	US-08-402-282-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-461-775-10
; Sequence 10, Application US/08461775
; Patent No. 5858773
GENERAL INFORMATION:
APPLICANT: MAZODIER, Philippe
APPLICANT: GUGLIEMI, Gerard
TITLE OF INVENTION: REGULATORY NUCLEOTIDE SEQUENCE OF THE
TITLE OF INVENTION: INITIATION OF TRANSCRIPTION
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,775
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/050,313
FILING DATE: 10-MAY-1993
APPLICATION NUMBER: FR 9011186
FILING DATE: 10-SEP-1990
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feuzy, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 010830-035
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1620 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1620
US-08-461-775-10

Query Match 3.8%; Score 47.6; DB 3; Length 1620;
 Best Local Similarity 45.7%; Pred. No. 0.0031;
 Matches 321; Conservative 0; Mismatches 369; Indels 12; Gaps 4;

120 CACCCGCGCGGTGCGGAGTAACTACCCCTGTCGATGAGTCTCTATCCAGGCGAG 179
 879 CACCTTCACCGGTGCGGAGTAACTACCCCTGTCGATGAGTCTCTATCCAGGCGAG 938
 180 TCGCGGAGAGGCGGTGCGGAGTAACTACCCCTGTCGATGAGTCTCTATCCAGGCGAG 239
 939 TCGGAGAGTCTGCGGAGTAACTACCCCTGTCGATGAGTCTCTATCCAGGCGAG 988
 240 CCGGAGAGTCTGCGGAGTAACTACCCCTGTCGATGAGTCTCTATCCAGGCGAG 239
 999 GAGGAGAGTCTGCGGAGTAACTACCCCTGTCGATGAGTCTCTATCCAGGCGAG 1058
 300 AGCGGAGAGTCTGCGGAGTAACTACCCCTGTCGATGAGTCTCTATCCAGGCGAG 359
 1059 CCGGAGAGTCTGCGGAGTAACTACCCCTGTCGATGAGTCTCTATCCAGGCGAG 1118
 360 C-GCGTACGCTGCTTGCAGCGATGCTGTTGATTCGCGACCACTCAATTGCTTACG 418
 1119 CCGGAGAGTCTGCGGAGTAACTACCCCTGTCGATGAGTCTCTATCCAGGCGAG 1178
 419 AGACGCTGCGGAGTAACTACCCCTGTCGATGAGTCTCTATCCAGGCGAG 476
 1179 GACCGCTGCGGAGTAACTACCCCTGTCGATGAGTCTCTATCCAGGCGAG 1238
 477 CCGTACGCTGCTTGCAGCGATGCTGTTGATTCGCGACCACTCAATTGCTTACG 533
 1239 CCGTACGCTGCTTGCAGCGATGCTGTTGATTCGCGACCACTCAATTGCTTACG 1298
 534 CCGTACGCTGCTTGCAGCGATGCTGTTGATTCGCGACCACTCAATTGCTTACG 591
 1299 CCGTACGCTGCTTGCAGCGATGCTGTTGATTCGCGACCACTCAATTGCTTACG 1358
 591 CCGTACGCTGCTTGCAGCGATGCTGTTGATTCGCGACCACTCAATTGCTTACG 647
 1359 CCGTACGCTGCTTGCAGCGATGCTGTTGATTCGCGACCACTCAATTGCTTACG 1418
 648 TACGCTGCTTGCAGCGATGCTGTTGATTCGCGACCACTCAATTGCTTACG 707
 1419 CCGTACGCTGCTTGCAGCGATGCTGTTGATTCGCGACCACTCAATTGCTTACG 1478
 708 GATCAGCTGCTTGCAGCGATGCTGTTGATTCGCGACCACTCAATTGCTTACG 767
 1479 CCGTACGCTGCTTGCAGCGATGCTGTTGATTCGCGACCACTCAATTGCTTACG 1538
 768 GATCAGCTGCTTGCAGCGATGCTGTTGATTCGCGACCACTCAATTGCTTACG 809
 1539 GATCAGCTGCTTGCAGCGATGCTGTTGATTCGCGACCACTCAATTGCTTACG 1580

RESULT 2
 US-08-258-261B-6
 Sequence 6, Application US/08258261B
 Patent No. 5639949

GENERAL INFORMATION:
 APPLICANT: Schupp, Thomas
 APPLICANT: Ligon, James M.
 APPLICANT: Beck, James Joseph
 APPLICANT: Hill, Dwight Steven
 APPLICANT: Ryals, John Andrew
 APPLICANT: Gaffney, Thomas Deane
 APPLICANT: Lam, Stephen Ting
 APPLICANT: Hammer, Phillip E.
 APPLICANT: Urnes, Scott Joseph
 TITLE OF INVENTION: Genes for the synthesis of
 TITLE OF INVENTION: antipathogenic substances
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESS: Ciba-Geigy Corporation

STREET: 7 Skyline Drive
 CITY: Hawthorne
 STATE: NY
 COUNTRY: USA
 ZIP: 10532
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION NUMBER: US/08/258, 261B
 FILING DATE: 08-JUN-1994
 CLASSIFICATION: 800
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 08/457, 205
 FILING DATE: 01-JUN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Elmer, James Scott
 REGISTRATION NUMBER: 36,129
 REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8614
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 28958 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOHETICAL: NO
 ANTI-SENSE: NO
 US-08-258-261B-6

Query Match 3.8%; Score 47.6; DB 1; Length 28958;
 Best Local Similarity 44.1%; Pred. No. 0.0076;
 Matches 286; Conservative 0; Mismatches 359; Indels 3; Gaps 2;

154 GCGCATGAGG1JCTCATCCAGGAGTGGCGGAGGCGCTCGCTATCCAGCGGAT 213
 25306 GCGCATGAGGAGTGGCGGAGTGGCGGAGGCGCTCGCTATCCAGCGGAT 25365
 214 TTCAAGGCGGAGTGGCGGAGTGGCGGAGGCGCTCGCTATCCAGCGGAT 273
 25366 TTCAAGGCGGAGTGGCGGAGTGGCGGAGGCGCTCGCTATCCAGCGGAT 25425
 274 TTATGCTAGGTCAGGAGTGGCGGAGTGGCGGAGGCGCTCGCTATCCAGCGGAT 333
 25426 TTATGCTAGGTCAGGAGTGGCGGAGTGGCGGAGGCGCTCGCTATCCAGCGGAT 25485
 334 ATCTGCTAGGTCAGGAGTGGCGGAGTGGCGGAGGCGCTCGCTATCCAGCGGAT 393
 25486 ATCTGCTAGGTCAGGAGTGGCGGAGTGGCGGAGGCGCTCGCTATCCAGCGGAT 25545
 394 TCGGCGAGTGGCGGAGTGGCGGAGTGGCGGAGGCGCTCGCTATCCAGCGGAT 453
 25546 TCGGCGAGTGGCGGAGTGGCGGAGTGGCGGAGGCGCTCGCTATCCAGCGGAT 25605
 454 CTTGCGGAGTGGCGGAGTGGCGGAGTGGCGGAGGCGCTCGCTATCCAGCGGAT 511
 25606 CTTGCGGAGTGGCGGAGTGGCGGAGTGGCGGAGGCGCTCGCTATCCAGCGGAT 25665
 512 TGATGCGAGTGGCGGAGTGGCGGAGTGGCGGAGGCGCTCGCTATCCAGCGGAT 571
 25666 TGATGCGAGTGGCGGAGTGGCGGAGTGGCGGAGGCGCTCGCTATCCAGCGGAT 25725
 572 CCGAGG-TGCTGCTAGTGGCGGAGTGGCGGAGTGGCGGAGGCGCTCGCTATCCAGCGGAT 630
 25726 CCGAGG-TGCTGCTAGTGGCGGAGTGGCGGAGTGGCGGAGGCGCTCGCTATCCAGCGGAT 25785
 631 GCGATGCGGAGTGGCGGAGTGGCGGAGTGGCGGAGGCGCTCGCTATCCAGCGGAT 690

DB 25786 GCGTGTGCGAGAGGCGCTCTCTCCCTCGAGAGAGCGGCGCCCGCATCGCCGCTCGGCAAGC 25845
 QY 691 GCCGAGTTCTGGCGGCGGATCCACACTCTACTCATCGGCGCTAGAGCTCGAGGTTGCC 750
 DB 25846 AAGGCCCTACACACCGCTCGCGGCAAGGCGGCGCATGCGCGCTCGAGCTCGGCGCTCC 25905
 QY 751 GTCGAAGCTGCGGAGCTGTGATTGGGCGCGCTCTGCTGTCAGGCGCC 798
 DB 25906 GACCTCAGACCTACTCCTCGCTCGGCGGAGAGGCGCTCTCATCGCC 25953

RESULT 3
 US-08-456-837-6
 ; Sequence 6, Application US/08456837
 ; Patent No. 5643774
 ; GENERAL INFORMATION:
 ; APPLICANT: Schupp, Thomas
 ; APPLICANT: Ligon, James M.
 ; APPLICANT: Beck, James Joseph
 ; APPLICANT: Hill, Dwight Steven
 ; APPLICANT: Ryals, John Andrew
 ; APPLICANT: Gaffney, Thomas Deane
 ; APPLICANT: Lam, Stephen Ting
 ; APPLICANT: Hammer, Phillip E.
 ; APPLICANT: Utnes, Scott Joseph
 ; TITLE OF INVENTION: Genes for the synthesis of
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Ciba-Geigy Corporation
 ; STREET: 7 Skyline Drive
 ; CITY: Hawthorne
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10532
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/456,837
 ; FILING DATE: 01-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/457,205
 ; FILING DATE: 01-JUN-1995
 ; APPLICATION NUMBER: 08/258,261
 ; FILING DATE: 08-JUN-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Elmer, James Scott
 ; REGISTRATION NUMBER: 36,129
 ; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 919-541-8614
 ; TELEFAX: 919-541-8689
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 28958 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; HYPOTHEICAL: NO
 ; ANTI-SENSE: NO
 ; US-08-456-837-6

Query Match 3.8%; Score 47.6; DB 1; Length 28958;
 Best Local Similarity 44.1%; Pred. No. 0.0076;
 Matches 286; Conservative 0; Mismatches 359; Indels 3; Gaps 2;

QY 154 GCGCATAGGTGCTCTCATCCAGGAGTGCAGAGGCGCTCGGCTATACACCGAGGAT 213
 DB 25306 GCCCAGACAGACCTCACCTCTGCGGATGTGCTATTGCTGCGCACACCGCGCCAC 25365
 QY 214 TTCAAGCGGCGAGCGCGGCAACTGTGTGCGGCGCGGAGCGAGTGTGTGGCGGAGT 273
 DB 25366 TTCCAGCAGCGCGCGGCTCTGTAGTACCCACACCGGAGAGTCTCTCTCGGCTGAC 25425
 QY 274 TTATGTCTCAAGTCAAGAAGACCGATAGCGCGGATATAGCGCGCTGTGACACAGG 333
 DB 25426 TCGCTCCCGCGGACAGACCGCGCGCGGAGCAGTCTCTGAGAGGAGGAGCAACGCG 25485
 QY 334 ATCTGTTCAGCTCTGCAATTGCGCGGCTGACGCTTGCACCGATGCTGTGGAT 393
 DB 25486 AAGCTGCTTGTCTTCTTCTGCGGCAAGCTTCCAGTGGAGAGATGCGCTCTGCTG 25545
 QY 394 TCCGAGCACACGTCATATTC :ACGAGACCGTCCAGACCGCGAGCGGCGACTACCCCTG 453
 DB 25546 CTCGACTCTCTGCGCGCTTCTCGGCTCAGTCTGACATGACGAGCGCGCTCGCTCT 25605
 QY 454 CTTGCCCCGATGAGCGAAGTGTGCGGCTGACATGCGCGCGGCTGTTGGCGTTACGAC -C 511
 DB 25606 CACGTGAGTGTGAGCTGTGCTGCGCTCTGCGCGGCGAGAGGCGCGCTCTCTGAC 25665
 QY 512 TGATCGAACCAGAGGCGCGCGGCTGTGATGAGCGCGGCGTCCCGCGCTGACCGG 571
 DB 25666 CGGTGCGAGCTGTGACACCGCG :CCTCTTCCGCTCATGCTCTCTCGGCGCGCTCTG 25725
 QY 572 CCGAGG-TGTGTGTATGCGGCGCGGCGCGGCTTACAAAGCAGCGCGCGATCGCCAC 630
 DB 25726 CGCTGCGTGTGAGCTGTGAGCGCGCGCGCTGCGCGGCGAGTGTGAGTGTGCGCG 25785
 QY 631 GGCATGCGGCGGCGGCTTACGCTTCTACATCAACATGCAAACTTGGCAACTGAC 690
 DB 25786 GCGTGTGCGAGGCGCTCTCTCGAGAGCGGCGCGCGCTGCGCGCTCTGACG 25845
 QY 691 GCCGAGTTCTGGCGGCGGATCCACACTGCTACTGAGCGCTAGAGCTCGAGGTTGCC 750
 DB 25846 AAGGCCCTACACACCGCTCGCGGCAAGGCGGCGCATGCGCGCTGAGCTCGGCGCTCC 25905
 QY 751 GTCGAAGCTGCGGAGCTGTGATTGGGCGCGCTCTGCTGTCAGGCGCC 798
 DB 25906 GACCTCAGACCTACTCCTCTGCGGCGGAGAGGCGCTCTCATCGCC 25953

RESULT 4
 US-08-457-342-6
 ; Sequence 6, Application US/08.57342
 ; Patent No. 5662898
 ; GENERAL INFORMATION:
 ; APPLICANT: Schupp, Thomas
 ; APPLICANT: Ligon, James M.
 ; APPLICANT: Beck, James Joseph
 ; APPLICANT: Hill, Dwight Steven
 ; APPLICANT: Ryals, John Andrew
 ; APPLICANT: Gaffney, Thomas Deane
 ; APPLICANT: Lam, Stephen Ting
 ; APPLICANT: Hammer, Phillip E.
 ; APPLICANT: Utnes, Scott Joseph
 ; TITLE OF INVENTION: Genes for the synthesis of
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Ciba-Geigy Corporation
 ; STREET: 7 Skyline Drive
 ; CITY: Hawthorne
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10532
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS


```

QY 274 TTATGCTCAAGTCAAGAACGATAGCGCGGATATACGGCCGCTTCGCACAGGGCAG 333
DB 25426 TCGCTCCCGACGACGACCGCGCCCGAGACCGCTCCCTCGAGGAGCGGAAAGCCAGGC 25485
QY 334 ATCTGTTCAAGTCTTGCAATTTGGCCGCGACGCTGCTTGCACGATGCTTTGGAT 393
DB 25486 AAGTCTGCTTCTGCTTCTTCCGCGCAAGGCTCGAGAGGAGGAGGAGGAGGAGGAGG 25545
QY 394 TCCGACACACGATGATGCTTACGATGACGATGACGATGACGATGACGATGACGATG 453
DB 25546 CTCGACTCTCTGCGCCGCTTCTCGCGCAGCTCGAAGCATGCGAGCGCGCTGCTCT 25605
QY 454 CTTCGCGGATGACGATGACGATGACGATGACGATGACGATGACGATGACGATGACG 511
DB 25606 CAGCTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 25665
QY 512 TGATGCGAACCCAAAGGGGCGCGGTGCTGATGAGGCGGCGCGCGCTCGAACCGG 571
DB 25666 CGCTCGACGCTGACGACGCGCGCGCTTTCGCTGATGATGCTCCCTGAGGCGCGCT 25725
QY 572 CCGAGC-TGCTGCTGATGCGCGCGCGCGCGCGCGCTGACGAGCGCGCGATGCGCAC 630
DB 25726 CGCTGCTGCGGCTGAGAGCGCGCGCGCGCTGCTGCGCGCACG CAGGCGGAGATGCG 3785
QY 631 GCGATGCGCGCGCGCTGATGCTGATGATGATGATGATGATGATGATGATGATGATG 690
DB 25786 GCGCTTCTGCGAGGCGCTCTCTCTCTGAGAGAGCGCGCGCGCGCTGCGCGAGC 25845
QY 691 GCGGAGTCTGCGCGCGCGGATGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 750
DB 25846 AAGCGCTGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 25905
QY 751 GTCAAGCTGCGCGCGCGCTGATGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 798
DB 25906 GACCTCCAGACTTACTGCTGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 25953

```

RESULT 6
US-08-458-076A-6
Sequence 6, Application US/08458076A
Patent No. 5698425

GENERAL INFORMATION:

APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Gaffney, Thomas Deane
APPLICANT: Lam, Stephen Ting
APPLICANT: Hammer, Phillip E.
APPLICANT: Ukner, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
NUMBER OF SEQUENCES: 22
TITLE OF INVENTION: antipathogenic substances

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA

SEQUENCE INFORMATION:

21P: 10532
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,076A
FILING DATE: 01-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205

FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-458-076A-6

Query Match 3.8% Score 47.6; DB 2; Length 28958;
Best Local Similarity 44.1%; Pred. No. 0.0076;
Matches 286; Conservative 0; Mismatches 359; Indels 3; Gaps 2;

```

QY 154 GCGGATGAGTGTCTCATCCAGCGAGGTCGCGGAGAGGCTCGGCTATCCAGCGCGGAT 213
DB 25306 GCGGATGAGTGTCTCATCCAGCGAGGTCGCGGAGAGGCTCGGCTATCCAGCGCGGAT 25365
QY 214 TTCAGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 273
DB 25366 TTCAGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 25425
QY 274 TTATGCT CAAGTCAAGAACGATAGCGCGGATGAGGCGCGCGCTGCGACAGGGCAG 333
DB 25426 TCGCTCCAGACTTACTGCTGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 25485
QY 334 ATCTGTTCAAGTCTTGCAATTTGGCCGCGACGCTGCTTGCACGATGCTTTGGAT 393
DB 25486 AAGTCTGCTTCTGCTTCTTCTGCGCAAGGCTCGCAGTGGAGGATGAGGCGCTGCTG 25545
QY 394 TCCGACACACGATGATGCTTACGATGACGATGACGATGACGATGACGATGACGATG 453
DB 25546 CTCGACTCTCTGCGCCGCTTCTCGCGCAGCTCGAAGCATGCGAGCGCGCTGCTCT 25605
QY 454 CTTCGCGGATGACGATGACGATGACGATGACGATGACGATGACGATGACGATGACG 511
DB 25606 CAGCTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 25665
QY 512 TGATGCGAACCCAAAGGGGCGCGGTGCTGATGAGGCGGCGCGCGCTCGAACCGG 571
DB 25666 CGCTCGACGCTGACGACGCGCGCGCTTTCGCTGATGATGCTCCCTGAGGCGCGCT 25725
QY 572 CCGAGC-TGCTGCTGATGCGCGCGCGCGCGCGCGCGCTGACGAGCGCGCGATGCGCAC 630
DB 25726 CGCTGCTGCGGCTGAGAGCGCGCGCGCGCGCTGCTGCGCGCACG CAGGCGGAGATGCG 25785
QY 631 GCGATGCGCGCGCGCTGATGCTGATGATGATGATGATGATGATGATGATGATGATG 690
DB 25786 GCGCTTCTGCGAGGCGCTCTCTCTGAGAGAGCGCGCGCGCGCGCGCGCGCGCG 25845
QY 691 GCGGAGTCTGCGCGCGCGGATGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 750
DB 25846 AAGCGCTGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 25905
QY 751 GTCAAGCTGCGCGCGCGCTGATGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 798
DB 25906 GACCTCCAGACTTACTGCTGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 25953

```

RESULT 7
US-08-764-233A-4

```

: Sequence 4, Application us/08764233A
: Patent No. 5716849
: GENERAL INFORMATION:
: APPLICANT: Ligon, James M.
: APPLICANT: Schupp, Thomas
: APPLICANT: Beck, James J.
: APPLICANT: Hill, Dwight S.
: APPLICANT: Neff, Snezana
: APPLICANT: Ryals, John A.
: TITLE OF INVENTION: Genes For The Biosynthesis Of Sorphen
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Ciba-Geigy Corporation
: STREET: 520 White Plains Road, P.O. Box 2005
: CITY: Tarrytown
: STATE: NY
: COUNTRY: USA
: ZIP: 10591
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/764,233A
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/729,214
: FILING DATE: 09-OCT-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/258,261
: FILING DATE: 08-JUN-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Meigs, J. Timothy
: REGISTRATION NUMBER: 38,241
: REFERENCE/DOCKET NUMBER: 1506/CIP6
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (919) 541-8587
: TELEFAX: (919) 541-8689
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 28958 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: ORIGINAL SOURCE:
: ORGANISM: Sorangium cellulosum
: IMMEDIATE SOURCE:
: CLONE: p98/1
: US-08-764-233A-4

```

Query Match 3.8%; Score 47.6; DB 2: Length 28958;
 Best Local Similarity 44.1%; Pred. No. 0.0076;
 Matches 286; Conservative 0; Mismatches 359; Indels 3; Gaps 2;

```

QY 154 GCGCATGAGGTGCTCATCCAGAGAGTGGCGGAGAGGGCTCTGACCGACGGAGAT 213
DB 25306 GCGCATGAGGTGCTCATCCAGAGAGTGGCGGAGAGGGCTCTGACCGACGGAGAT 25365
QY 214 TTCAGAGCGGACGCGCAACTGTGCGGACCGCGACAGGATGTGGGCGGAGCTGAT 273
DB 25366 TTCAGAGCGGACGCGCGCTGTGCGGACCGCGACAGGATGTGGGCGGAGCTGAT 25425
QY 274 TTATGCTCAAGTCAAGAACGATAGCGGCGGATAGCGGCGCTGGGACGACGGGAG 333
DB 25426 TCCTCGCCGACGAGAACCGCGGCGGAGACCGCTCTGCGAGGAGCGGAGACGAGG 25485
QY 334 ATCTGTTCACGTTCTGATGAGTTGGCGCGCTGACGTTGACCGATGCGTTGTTGAT 393
DB 25486 AAGCTGCTTTCGCTTCTGCGGACGAGGCTCGCAGTGGAGAGGATGGCCCTCTGCTG 25545

```

```

QY 394 TCCGGCACCACGTCATTCCTTACGAGACCGTCCAGACCGCCGAGCGGACCTACCTCTG 453
DB 25546 CTCGACTCCTCGCCCGCTTCTCCGCTCAGCTCGAAGATGAGACGCGCCCTGCTCT 25605
QY 454 CTTCGCCGATGAGGAGTGGCGGATGAGTGGCGGACGAGTGGCGGCTTACAC--C 511
DB 25606 CAGCTGAGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 25665
QY 512 TGTGCGAACCCAGAGGCGCGCGGTGCTGATGAGCGGCGGCGGCGGCTGCAACCGG 571
DB 25666 CGGCTGAGCTGTGATGAGCGCGCGCGGTGCTGATGAGCGGCGGCGGCGGCTGCA 25725
QY 572 CCGACG-TGTGTTATTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 630
DB 25726 CGCTGCTGCGCGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 25785
QY 631 GCGATGGGCGGACCGCTTACGATGATGAGATCAACATGACAGAACTTGGCACTCGAC 690
DB 25786 GCCTGTGTCGACAGCGCGCTTCTGCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCG 25845
QY 691 GCGGATTCGCGCGCGCGGATCCACTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 750
DB 25846 AAGCGCTGACACCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 25905
QY 751 GTCAACGTGCGCGCGCTGATGATGAGGCGCGCTGCTGCTGCGGCGCGCGCGCGCG 798
DB 25906 GACCTCGACACTACCTCGCTGCTGCGGCGGACAGGCTCTCATCCGC 25953

```

RESULT 8
 US-08-457-335A-6
 : Sequence 6, Application us/08457335A
 : Patent No. 5723759
 : GENERAL INFORMATION:
 : APPLICANT: Schupp, Thomas
 : APPLICANT: Ligon, James M.
 : APPLICANT: Beck, James Joseph
 : APPLICANT: Hill, Dwight Steven
 : APPLICANT: Ryals, John Andrew
 : APPLICANT: Gaffney, Thomas Deane
 : APPLICANT: Lam, Stephen Tung
 : APPLICANT: Hammer, Phillip E.
 : APPLICANT: Umes, Scott Joseph
 : TITLE OF INVENTION: Genes for the synthesis of
 : TITLE OF INVENTION: antipathogenic substances
 : NUMBER OF SEQUENCES: 22
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Ciba-Geigy Corporation
 : STREET: 7 Skyline Drive
 : CITY: Hawthorne
 : STATE: NY
 : COUNTRY: USA
 : ZIP: 10532
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/457,335A
 : FILING DATE: 01-JUN-1995
 : CLASSIFICATION: 800
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/457,205
 : FILING DATE: 01-JUN-1995
 : APPLICATION NUMBER: 08/258,261
 : FILING DATE: 08-JUN-1994
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Elmer, James Scott
 : REGISTRATION NUMBER: 36,129
 : REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
 : TELECOMMUNICATION INFORMATION:

```

? TELEPHONE: 919-541-8614
? TELEFAX: 919-541-8689
? INFORMATION FOR SEQ ID NO: 6:
?
? SEQUENCE CHARACTERISTICS:
?
? LENGTH: 2895 base pairs
?
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? HYDROTHERICAL: NO
?
? ANTI-SENSE: NO
US-08-457-335A-6

```

Query Match	3.8%;	Score 47.6;	DB 2;	Length 28358;
Best Local Similarity	44.1%;	Pred. No. 0.0076;		
Matches 286;	Conservative	0;	Mismatches 359;	Indels 3;
				Gaps 2

OY	154	GGCATGAGAGTCTCATCCAGGACAGGTGGCGGAGAGGGCTCGGATATCCACCAAGCGAT	213
Db	25306	GGCCACGACGACCTCAACCTTCGGCGGATGAGGCGCTATTCGCTGGCCACACACCGCGGCCAC	25365
OY	214	TTCAAGGGGGGAGGGCGGCACTGGTCGGACCGCGGACCAAGGTGTGGCGACCGCTGAT	273
Db	25366	TTCGAGCACCGGGCGCTCTCGTAGGCCCAACCGGACGAGACTCTCTCCGCGCTCGAC	25425
OY	274	TTATTGCTCAAGGTCAAGAACCAGATAGGGGGGGAATAGGGCGGCTGGACACAGGGGAG	333
Db	25426	TGCGTCGGCCCAAGACAGCCGCCGCCGACACCGTCTCGAGCGAGCGGAAGCCAGCGC	25485
OY	334	ATCTGTTCACGCTTCTTCGATTGGCGCGGTACAGGCTTTCGACAGATCGTTGTTGAT	393
Db	25486	AAGCTGCTTCTGTCTTCTTCTGGGCAAGGCTCGCAATGGGAAGGAATGGCCCTCTGCTG	25545
OY	394	TCCGGACACGACTCAATTGCTCTACGAGACGTTCCAGACCCGCGAGCGGCACATCCCTG	453
Db	25546	CTCGACTCTCCGCCGCTTTCGGCGCTCAGCTCGAAGCATTCGAGAGCGCGCTCGCTCT	25605
OY	454	CTTGCCCCGATAGGGAATTCGCGCGTGCACATCGCGCCAGTTGGCGCTTACCAAC--C	511
Db	25606	CACGTGATGAGAGCTCTGTCGCGCTTCGCGCGGAGAGAGGGCCCCCTCTCCTCGAC	25665
OY	512	TCATGCGACCCCAAGGGGGGCGCGGTGTCTGATGGCGGGGTGCCCGGCTGGAACCG	571
Db	25666	CGGTGACAGCTGTAAAGCCCGCCCTCTTTGGCGTATGATGTCTCCCTGGCGGCTCTGG	25725
OY	572	CCGAGC-TCGTGATGCGCGCGCGGACCGCGGCTACAGCGACCGCATCGCCAAAC	630
Db	25726	CCGTGCTCGGGGTAGAGCCGCGCGCGCTGTGCGGCACATGAGGGCAGATCGCGGCC	25785
OY	631	GGCATGGGGCGACCTTAAGCTTCTAGACATTCACATGACAAATTGGCAACTCGAC	690
Db	25786	GGCTTGTGCGAGGGCGCTCTTCCTCTGAGAGAGCGGCCCGCATCGCGGCTGGCGACG	25844
OY	691	GCCGATTCGCGGCGCGGATTCACATCTCCTACTCATCGGCTACGAGCTGAGGGTGGC	750
Db	25846	AAAGGCTCACACAGCTGCGCGGCAACGGGGGCATGGCGCGCTGAGGCTCGGCTC	25906
OY	751	GTCMAACGTCCGACCTGGTGAATGGGGGCGTCTCTGTGCCAGCGCC	798
Db	25906	GACCTTCACACCTCACTCGCTCTCCCTGGGGGAGAGGCTCTC-ATCGCC	25953

RESULT 9
US-08-729-214-6
Sequence 6, Application US/08729214
Patent No. 5617502
GENERAL INFORMATION:
APPLICANT: Ligon, James M.
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Hammer, Phillip E.
APPLICANT: van Pee, Karl-Heinz

```

1  APPLICANT: Kliner, Sabine
2  TITLE OF INVENTION: Genes for the synthesis of
3  TITLE OF INVENTION: antipathogenic substances
4  NUMBER OF SEQUENCES: 27
5  CORRESPONDENCE ADDRESS:
6  ADDRESSEE: Ciba-Geigy Corporation
7  STREET: 520 White Plains Road
8  CITY: Tarrytown
9  STATE: NY
10 COUNTRY: USA
11 ZIP: 10591
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE: Floppy disk
14 COMPUTER: IBM PC compatible
15 OPERATING SYSTEM: PC-DOS/MS-DOS
16 SOFTWARE: PatentIn Release #1.0, Version #1.25
17 CURRENT APPLICATION DATA:
18 APPLICATION NUMBER: US/08/729,214
19 FILING DATE: TBA
20 CLASSIFICATION: 435
21 ATTORNEY/AGENT INFORMATION:
22 NAME: Meigs, J. Timothy
23 REGISTRATION NUMBER: 38,241
24 REFERENCE/DOCKET NUMBER: CGC 1506/CIPS
25 TELECOMMUNICATION INFORMATION:
26 TELEPHONE: 919-541-8587
27 TELEFAX: 919-541-8699
28 INFORMATION FOR SEQ ID NO: 6:
29 SEQUENCE CHARACTERISTICS:
30 LENGTH: 28958 base pairs
31 TYPE: nucleic acid
32 STRANDEDNESS: single
33 TOPOLOGY: linear
34 MOLECULE TYPE: DNA (genomic)
35 HYPOTHEICAL: NO
36 ANTI-SENSE: NO
37
38 US-08-729-214-6

```

Query Match	3.6%;	Score 47.6;	DB 2;	Length 28958;
Best Local Similarity	44.1%;	Pred. No. 0.0076;		
Matches 286;	Conservative	0;	Mismatches 359;	Indels 3;
				Gaps 2

QY	154	GGCAGTAGAGTCTCTATCCAGGACAGATGGCCGAGAGAGGCTCGGCTATCACAGCAGGGAT	213
Db	25306	GCCACGACGACACTCACCTCGCGGATGTGGCTTATCGCTTGCCATGCCACACCCGCGCCAC	25365
QY	214	TTCAAGGCGGACGCCUCCAMTGGTGTCTTCTTCCGACAGAGTGTGGCCGACGCTGAT	273
Db	25366	TTGACACACCGGGCCCTCTTCCTATCTCAACACCGGACGACACTCTCTCTCGCGGCTGCAC	25425
QY	274	TTATTGCTCAAGGTCAAGAACCGATACGGCGGAATACGGCGCCCTCGACACGGGCGAG	333
Db	25426	TGCGTGGCCAGACAAAGCCGCCCGGACACCGTCTCGACAGGAGCGAAGCCACGGC	25485
QY	334	AICTTGTTCACGTTCTTGATTTGGCCCGCTACGTCGTTCGACCGATCCGTTGTGGAT	393
Db	25486	AAGCTCGTCTTGCTCTTCTTCCTGGGCAAGGCTGCACTGGGAAGGATGCGCTCTCGCTG	25545
QY	394	TCCGGACACAGCAATTCCTCTACGAGACCGTCCAGACCGCGGACGAGCGGCACTTACCCTG	453
Db	25546	CTGACACTCTTCGCCGCTCTTCGCGCTACGCTCGAAGCATGGAGAGCGCGCTCGCTCT	25605
QY	454	CTTGCCCCGATGACGAAAGTCGCGGCTGACATCGCCGCCCAAGATTGGCGCTTACGAC-C	511
Db	25606	CAGCTGAGATGGAGCCTGCTCGCGGCTCTGCGCGGAGACGAGAGGGGCCCCCTCCCTCGAC	25665
QY	512	TGATGCGAAACCCAAAGGGGCGCGGGTGTCTGTATGGGCGGGGTGCCCGGCTGCAACGG	571
Db	25666	CGGCTGACGATGATACGCCGCCCTCTTTCGCGATGATGTCTCCCTCGGCGGCGCTCTTG	25725
QY	572	CGGAG-TCGTGTGATCGGCGCGGCGACCGCGGCTAACACGACCGCGCATCGCCAAC	630

Db 25726 GCGTCGCTGGCGATGAGACCCGCCGCTCGTCGCGCACAGTCAAGGGGAGATGCGCGCC 25785
 QY 631 GCGATGGCGCGGACCGTTAGCGTTCTAGACATCAACATGCAAACTTGGCACTGCGAC 690
 Db 25786 GCGTCGTCGAGGCGGCTCTCTCTCTCGAGAGAGCGGGCCGCGCTGCGCGACG 25845
 QY 691 GCCGAGTCTGCGGGCGGATCCACACTGCTACTGCGCTACGAGCTGAGAGGTGCG 750
 Db 25846 AAGGCGCTCACCACCGTCCGCGGCAAGCGGCGCATGCGCGCTGCGAGCTCGCGCTCC 25905
 QY 751 GTCAAACGCTGCCGACGTGTGATTGGGCGCTCTGTGTGCGAGGCGCC 798
 Db 25906 GACGTCGACACTGACTGCTGCTCCCTGGGGCGCAGAGGCTCTCATGCGCC 25953

RESULT 10
 US-08-764-233A-1
 : Sequence 1, Application US/08764233A
 : Patent No. 5716849
 :
 : GENERAL INFORMATION:
 : APPLICANT: Ligon, James M.
 : APPLICANT: Schupp, Thomas
 : APPLICANT: Beck, James J.
 : APPLICANT: Hill, Dwight S.
 : APPLICANT: Neff, Snezana
 : APPLICANT: Ryals, John A.
 : TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen
 : NUMBER OF SEQUENCES: 10
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Ciba-Geigy Corporation
 : STREET: 520 White Plains Road, P.O. Box 2005
 : CITY: Tarrytown
 : STATE: NY
 : COUNTRY: USA
 : ZIP: 10591
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent Release #1.0, Version #1.30
 :
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/764,233A
 : FILING DATE:
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/729,214
 : FILING DATE: 09-OCT-1996
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/258,261
 : FILING DATE: 08-JUN-1994
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Meigs, J. Timothy
 : REGISTRATION NUMBER: 38,241
 : REFERENCE/DOCKET NUMBER: 1506/CIP6
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (919) 541-8587
 : TELEFAX: (919) 541-8689
 : INFORMATION FOR SEQ ID NO: 1:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 49377 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: DNA (genomic)
 : ORIGINAL SOURCE:
 : ORGANISM: Sorangium cellulosum
 : IMMEDIATE SOURCE:
 : CLONE: p98/1, pUL3, and pYKM15
 : FEATURE:
 : NAME/KEY: misc_feature
 : LOCATION: 383..760
 : OTHER INFORMATION: /product= "SorB"
 : OTHER INFORMATION: /note= "This gene encodes a protein that is highly homologous

OTHER INFORMATION: the reductase domains of type I PKs such as eryA from
 OTHER INFORMATION: Saccharopolyspora erythraea."
 :
 : FEATURE:
 : NAME/KEY: misc_feature
 : LOCATION: 927..19874
 : OTHER INFORMATION: /product= "SorA"
 : OTHER INFORMATION: /note= "Gene product is highly homologous to type I PKs t
 : OTHER INFORMATION: are known to be involved in the synthesis of polyketide
 : OTHER INFORMATION: compounds."
 :
 : FEATURE:
 : NAME/KEY: misc_feature
 : LOCATION: 942..7115
 : OTHER INFORMATION: /product= "Module 1 of SorA"
 :
 : FEATURE:
 : NAME/KEY: misc_feature
 : LOCATION: 7203..12884
 : OTHER INFORMATION: /product= "Module 2 of SorA"
 :
 : FEATURE:
 : NAME/KEY: misc_feature
 : LOCATION: 13455..19616
 : OTHER INFORMATION: /product= "Module 3 of SorA"
 :
 : FEATURE:
 : NAME/KEY: misc_feature
 : LOCATION: 19871..46318
 : OTHER INFORMATION: /product= "SorB"
 : OTHER INFORMATION: /note= "Gene product is highly homologous to type I PKs ge
 :
 : FEATURE:
 : NAME/KEY: misc_feature
 : LOCATION: 19870..24556
 : OTHER INFORMATION: /product= "Module 1 of SorB"
 :
 : FEATURE:
 : NAME/KEY: misc_feature
 : LOCATION: 24638..30820
 : OTHER INFORMATION: /product= "Module 2 of SorB"
 :
 : FEATURE:
 : NAME/KEY: misc_feature
 : LOCATION: 30881..35446
 : OTHER INFORMATION: /product= "Module 3 of SorB"
 :
 : FEATURE:
 : NAME/KEY: misc_feature
 : LOCATION: 35528..40111
 : OTHER INFORMATION: /product= "Module 4 of SorB"
 :
 : FEATURE:
 : NAME/KEY: misc_feature
 : LOCATION: 40190..46318
 : OTHER INFORMATION: /product= "Module 5 of SorB"
 :
 : FEATURE:
 : NAME/KEY: misc_feature
 : LOCATION: 46851..47891
 : OTHER INFORMATION: /product= "SorM"
 : OTHER INFORMATION: /note= "The protein encoded by the sorM gene is highly
 : OTHER INFORMATION: homologous to the methyltransferase from Streptomyces
 : OTHER INFORMATION: hygroscopicus that is involved in the synthesis of the
 : OTHER INFORMATION: polyketide rapamycin."
 :
 : US-08-764-233A-1

Query Match 3.88; Score 47.6; DB 2; Length 49377;
 Best Local Similarity 44.1%; Pred. No. 0.0089;
 Matches 286; Conservative 0; Mismatches 359; Indels 3; Gaps 2;

QY 154 GGCATGAGTGCATTCATCCAGGAGGTCCGAGAGAGGCTCGCTACACCGAGCGGAT 213
 Db 41609 GCCCAGACGACCTCCTCGGATGTGCTATTCGTCGACACCGCGGCCAC 41668
 QY 214 TTCAGCGGCGGCGGCGCAACTGTCGCGACCGCGACCGAGTGTGGCGAGCTGAT 273
 Db 41669 TTCAGACCGCGGCGGCTCTCTGTCGACCAACCGCGACGAGCTCTCTCGCGCTGAC 41728
 QY 274 TTATGCTCAAGTCAAGAAGACGATAGCGCGGATATAGCGCGCTGCGACAGGGCAG 333
 Db 41729 TCGCTGCGCCGAGCAGACCGCGGCCCGGAGACGCTCTCGAGCGAGCGAGGACGACGCGC 41788

QY 334 ATCTGTACGTTCTTCATTTGGCCGCGTCAAGTCTTGACACGATGTTGGAT 393
 DB 41789 AAGCTGCTCTTCTTCTTCTTGCGCAAGGCTCGAGTGGGAAGGGATGGCCCTGCTG 41848
 QY 394 TCCGGACACAGTCAATTGCTTACAGAGCCGTCCAGACCCCGGCGCACTACCCCTG 453
 DB 41849 CTCGACTCTCGCCCGCTTCCGCGCTCAGCTCAACCATGAGCGCGCTGCTCTCT 41908
 QY 454 CTGCCCCCGGAGGAGGAGCGCGCTGCACTGCGCCGCAAGTGGCGCTTACAC -C 511
 DB 41909 CAGTCGAGTGGAGCTGCTGCGCTGCTGCGCCGCAAGTGGCGCTTACAC 41968
 QY 512 TGATCGCAACCAAGGCGCGCGCTGCTGATGCGCGGGTCCCGCGCTCAACCG 571
 DB 41969 CGGCTGACAGCTGTCAGAGCCCGCCCTTTCGCTCATGCTCTCCGCGCGCTGCG 42028
 QY 572 CCGAGC-TCTGTGTGATCGCGCGCGCGCGCGCTCAACAGAGCCCGATGCGCAC 630
 DB 42029 CGCTGCTCGCGCTGAGAGCCCGCGCGCTGCTGCGCACTGAGCGAGATGCGCGC 42088
 QY 631 GGCATGGCGCGACCTTACGCTTACATCAACATCAACATTCGCGCACTGAC 690
 DB 42089 GCTTCTGCTGAGCGCTCTCTCTGAGAGAGCGCGCGCGCTGCGCGCTGCGAGC 42148
 QY 691 GCCGAGTTTGGCGCGCGATCCACTGCTACTATCGCGCTTACGAGCTGAGGCTGCC 750
 DB 42149 AAACGCTACACACGCTGCGCGCGCAAGGCGCATGCGCGCTGAGCTGCGCGCTCC 42208
 QY 751 GTCAAGGTGCGACCTGCTGATGCGCGCGCTGCTGCGCGCGCGCGCGCGCGCG 798
 DB 42209 GACCTCAGACT 42256

RESULT 11

US-08-816-105A-2

Sequence 2, Application US/08816105A

Patent No. 5989882

GENERAL INFORMATION:

APPLICANT: Crieger, Anneliese; Dellweg, Hans-Georg; Lenz,
 APPLICANT: Jurgen; Schröder, Werner; Pape, Hermann;
 APPLICANT: Goeke, Klaus; Schaper, Beate; Hemker, Michael;
 APPLICANT: Piepersberg, Wolfgang; Distler, Jürgen;
 APPLICANT: Strittmann, Ansgar
 TITLE OF INVENTION: PROCESSES FOR PREPARING ACAROTOSYL
 TITLE OF INVENTION: TRANSFERASE AND FOR USING IT IN THE
 TITLE OF INVENTION: CONVERSION OF ACAROSE HOMOLOGUES
 TITLE OF INVENTION: INTO ACAROSE, FOR THE PREPARATION
 TITLE OF INVENTION: OF ACAROSE HOMOLOGUES
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Sprung Kramer Schaefer & Briscoe
 STREET: 660 White Plains Road
 CITY: Tarrytown
 STATE: New York
 COUNTRY: USA

ZIP: 10591-5144

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MB storage

COMPUTER: Apple Macintosh 6500

OPERATING SYSTEM: System 7.5

SOFTWARE: WordPerfect 3.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/816, 105A

FILING DATE: 14-MAR-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 19625269.5

FILING DATE: 25-JUN-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 19611252.4

FILING DATE: 22-MAR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Kurt G. Briscoe

REGISTRATION NUMBER: 33,141
 REFERENCE/DOCKET NUMBER: Bayer 9814-KGB
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (914) 332-1700
 TELEFAX: (914) 332-1844
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2582 nucleotides
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear

US-08-816-105A-2

Query Match 3.7%; Score 46; DB 4; Length 2582;
 Best Local Similarity 47.3%; Pred. No. 0.009;
 Matches 139; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

QY 359 CCGCGTACGTGCTTGCACCGATGCGTTGATTCGCGCACCGTCAATTGCTTACG 418
 DB 1576 CCGCGTACGTGCTTGCACCGATGCGTTGATTCGCGCACCGTCAATTGCTTACG 1635
 QY 419 AGACGCTGACAGCG 478
 DB 1636 TGACCGTGAAGAGCG 1695
 QY 479 GTCGACTGCG 538
 DB 1696 GCGAGTGTGCG 1755
 QY 539 TGTGATGCG 598
 DB 1756 TCGGCGCGGTGTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1815
 QY 599 CCGCGCGCTGACAGCG 652
 DB 1816 CCGCGCGACCGCTGCGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1869

RESULT 12

US-08-403-852D-3

Sequence 3, Application US/08403852D

Patent No. 5891695

GENERAL INFORMATION:

APPLICANT: Blanc, Veronique
 APPLICANT: Blanche, Francis
 APPLICANT: Crouzet, Joel
 APPLICANT: Jacques, Nathalie
 APPLICANT: Lacroix, Patricia
 APPLICANT: Thibaut, Denis
 APPLICANT: Zagorec, Monique
 APPLICANT: Debussche, Laurent
 APPLICANT: De Crey-Lagard, Valerie
 TITLE OF INVENTION: Polypeptides Involved In The
 TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences
 TITLE OF INVENTION: Coding For These Polypeptides And Their Use
 NUMBER OF SEQUENCES: 43
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Flinnegan, Henderson, Farabow, Garrett & Dunner
 STREET: 1300 I Street, N.W., Suite 700
 CITY: Washington
 STATE: D.C.

COUNTRY: USA

ZIP: 20005-3115

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/403, 852D

FILING DATE: 10-MAY-1995

PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: PCT/FR 93/00923
FILING DATE: 25-SEP-1993
PRIOR APPLICATION DATA: FR 92/11441
APPLICATION NUMBER: FR 92/11441
FILING DATE: 25-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03806.0054-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 833 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: S.pristinaespiralis
FEATURE:
NAME/KEY: CDS
LOCATION: 1..833
US-08-403-852D-3

```

```

Query Match          3.6%; Score 45.4; DB 3; Length 833;
Best Local Similarity 48.0%; Pred. No. 0.009;
Matches 130; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

```

```

QY 549 CGGGGTCCCGGCTCGAAGCGGCGGAGCTGTGATCGGCGCGGCGGCTA 608
    || || || || || || || || || || || || || || || || || ||
DB 327 CGAGGGCGCGGACCGGCGAATCGAGCTGTCCGGGCGCTGTGGACAGTCA 386
QY 609 CAACGAGCGCGCATCGCCAAAGCGGCGGCGGCGGCTT 367TTAGACATCA 668
    || || || || || || || || || || || || || || || || || ||
DB 387 CGAGCGCTGTGTCACGACGCGCGGCGGCTGTACTGCGGCGGCGGCTCA 446
QY 669 CGACAACTTCGGAATCGAGCGGCGGAGTTCTGGGCGGATCCAGCTGCTATC 728
    || || || || || || || || || || || || || || || || || ||
DB 447 ACTGACCGACCGGCGGCGGCGGAGCTGTGAGTGCGGCGGCGGCGGCGGCG 506
QY 729 GGCCTACGAGCTCGAGGCGGCGGCTCAAAAGCTGCGGCTGTGATTGGGCGGCTGTGT 788
    || || || || || || || || || || || || || || || || || ||
DB 507 CGAGGGCGGACCGGCTGTGCTGCGGCTGACCGGCGGCGGCGGCGGCGGCGG 566
QY 789 GCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 819
    || || || || || || || || || || || || || || || || || ||
DB 567 CCGTCTCGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 597

```

```

RESULT 13
US-08-403-852D-1
; Sequence 1, Application US/08403852D
; Patent No. 5891695
; GENERAL INFORMATION:
; APPLICANT: Blanc, Veronique
; APPLICANT: Blanchet, Francis
; APPLICANT: Crouzet, Joel
; APPLICANT: Jacques, Nathalie
; APPLICANT: Lacroix, Patricia
; APPLICANT: Thibaut, Denis
; APPLICANT: Zagorec, Monique
; APPLICANT: Debussche, Laurent
; APPLICANT: De Crey-Lagard, Valerie
; TITLE OF INVENTION: Polypeptides Involved In The
; TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences
; NUMBER OF SEQUENCES: 43
; CODING FOR THESE POLYPEPTIDES AND THEIR USE
; CORRESPONDENCE ADDRESS:

```

```

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/403,852D
FILING DATE: 10-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR 93/00923
FILING DATE: 25-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/11441
FILING DATE: 25-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03806.0054-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5392 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: S.pristinaespiralis
US-08-403-852D-1

```

```

Query Match          3.6%; Score 45.4; DB 3; Length 5392;
Best Local Similarity 48.0%; Pred. No. 0.016;
Matches 130; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

```

```

QY 549 CGGGGTCCCGGCTCGAAGCGGCGGAGCTGTGATCGGCGGCGGCGGCTA 608
    || || || || || || || || || || || || || || || || || ||
DB 3018 CGAGGGCGGCGGCGGCGGAGCTGTGAGTGCGGCGGCGGCGGCGGCGGCG 3077
QY 609 CAACGAGCGCGCATCGCCAAAGCGGCGGCGGCGGCGGCTTACGATTAGACATCA 668
    || || || || || || || || || || || || || || || || || ||
DB 3078 CGAGCGCTGTGTCACGACGCGCGGCGGCGGCTGTACTGCGGCGGCGGCGGCG 3137
QY 669 CGACAACTTCGGAATCGAGCGGCGGAGTTCTGGGCGGAGTCCAGCTGCTATC 728
    || || || || || || || || || || || || || || || || || ||
DB 3138 ACTGACCGACCGGCGGCGGCGGAGCTGTGAGTGCGGCGGCGGCGGCGGCGG 3197
QY 729 GGCCTACGAGCTCGAGGCGGCGGCTCAAAAGCTGCGGCGGCGGCTGTGATTGGGCGGCTGTGT 788
    || || || || || || || || || || || || || || || || || ||
DB 3198 CGAGGGCGGACCGGCTGTGCTGCGGCTGACCGGCGGCGGCGGCGGCGGCGG 3257
QY 789 GCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 819
    || || || || || || || || || || || || || || || || || ||
DB 3258 CCGTCTCGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3288

```

```

RESULT 14
US-08-403-852D-6
; Sequence 6, Application US/08403852D
; Patent No. 5891695
; GENERAL INFORMATION:
; APPLICANT: Blanc, Veronique
; APPLICANT: Blanchet, Francis

```


Db 1310 CCCGAGGCGCTCTCTACACCGAGCCGTCGTCTGCGGCTTCATGTCGCGCGCGAGCA 1369
 QY 651 GGTTCCTAGACATCAACATGCAAACTTGGCACTGAGCCGAGTCTGCGCGCGAT 710
 Db 1370 CGGCGTCGACGTGACCAACAGCTATTACACCGACCCGTCGTAATTCACTGCAAGCA 1429
 QY 711 CCACACTCGCTACTCATCGGCTACGAGCTCGAGGGTCCGTCAAACGTGCGGACCTGGT 770
 Db 1430 CGACCCCGACGAGAGGCGCTCGTCGAGGCGCCGTCCTGCGGGCTCCCGGTACGCGGAGAA 1489
 QY 771 GATTGGGCGCTCTGCTGTCGCAAGCGCC 798
 Db 1490 GAAGGCGCGCTCAACGTCGCGCGCGCC 1517

Search completed: June 22, 2000, 15:01:15
 Job time: 17553 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 12:07:07 ; Search time 5541.94 Seconds
(without alignments)
910.559 Million cell updates/sec

Title: us-09-362-485-2

Perfect score: 1245
Sequence: 1 ATCTGCAGATTATTCGAC.....AAGCGATGATCGCGCGG 1245

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 202611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:*

1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: gb_est33:*
53: gb_est34:*
54: gb_est35:*
55: gb_est36:*
56: gb_est37:*
57: gb_est38:*
58: gb_est39:*
59: gb_est40:*
60: gb_est41:*
61: gb_est42:*
62: gb_est43:*
63: gb_est44:*
64: gb_est45:*
65: gb_est46:*
66: gb_est47:*
67: gb_est48:*
68: gb_est49:*
69: gb_est50:*
70: gb_est51:*
71: gb_est52:*
72: gb_est53:*
73: gb_est54:*
74: gb_est55:*
75: gb_est56:*
76: gb_est57:*
77: gb_est58:*
78: gb_est59:*
79: gb_est60:*
80: gb_est61:*
81: gb_est62:*
82: gb_est63:*
83: gb_est64:*
84: gb_est65:*
85: gb_est66:*
86: gb_est67:*
87: gb_est68:*
88: gb_est69:*
89: gb_est70:*
90: gb_est71:*
91: gb_est72:*
92: gb_est73:*
93: gb_est74:*
94: gb_est75:*
95: gb_est76:*
96: gb_est77:*
97: gb_est78:*
98: gb_est79:*
99: gb_est80:*
100: gb_est81:*
101: gb_est82:*
102: gb_est83:*
103: gb_est84:*
104: gb_est85:*
105: gb_est86:*
106: gb_est87:*
107: gb_est88:*
108: gb_est89:*
109: gb_est90:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result

%
Query

SUMMARIES

AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

BP91 91006 EVRI cedex - FRANCE (E-mail : secretegenscope.cns.it)
 - Web : www.genoscope.cns.it)
 determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see <http://www.fruitfly.org> The BDGP Drosophila
 melanogaster BAC library was prepared by Kaitoyo Osoegawa and
 Aaron Mamoser in Pieter de Jong's laboratory in the Department of
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
 NY. The library is named RPc1-98 and was constructed by partial
 EcoRI digestion of Drosophila DNA provided by the BDGP from the
 isogenic strain y2; cn bw sp, the same strain used for the BDGP's
 p1 and EST libraries. A more detailed description of the library
 and how to order individual BAC clones, the entire library, or
 filters for hybridization from the BACPAC Resource Center can be
 found at http://dbcpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES	source	Location/Qualifiers
	1. .925	
	/organism="Drosophila melanogaster"	
	/db_xref="taxon:7227"	
	/clone_1lb="RPC1-98"	
	/clone="BACR19D16"	
	/note="end : TET3"	
BASE COUNT	120 a	61 g 172 t 511 others
ORIGIN	61 c	

Query Match	4.4%	Score 55.2	DB 82	Length 925
Best Local Similarity	11.8%	Pred. No. 0.15		
Matches 41	Conservative 170	Mismatches 137	Indels 0	Gaps 0
QY 431	CCGCCGACGGCGCTACCCCTGCTTGCCCGCATAGCGCAAGTCGCCGGTGCATCGCCG			490
Db 577	SCSCSSSCSCCBCCCCCSCSSCYCCSSSSSSKSCSTSBSCSCCCKSVCTCSCSSSS			636
QY 491	CCGAGTGGGCCCTTACGACCTGATGGCAACCAAGGGGGCGCGGTGCTGATAGGGCG			550
Db 637	CSSSSSTSSSSSTSSSTSKSSSSGSSSSSTTTSKTSASGSGWSAGGSGSTGSTS			696
QY 551	GGGTCCCGCGCGCTCGAACCCGCCGACGTCGTGATCGCGCCGACACCGCCGCTACA			610
Db 697	SSSSSSSTSTSSSVSGKSTBSGSSBSSSGSSSSSTSTSBSCOTSTSSSSSSSVSS			756
QY 611	ACGCAGCCCGCATCCGCCAACGGCATTGGGGCGCACCGTTACGTTTACGATCAACATCG			670
Db 757	TCSCCTCCSCSYSTSSSTSSSTSMGTSISGSSSSVGTSSSDSTSTCCSCCYMCTCS			816
QY 671	ACAACCTTCGCGAATCGACGCCGAGATTGTGGCGCGCGATCCACATCGCTACTATCGG			730
Db 817	TYBMCYITSTGCGSSSSSGKGGYVTKCGGGGSSSTNMBETSSACSSSSSSCGSSSVS			876
QY 731	CCTACGAGCTCGAGGCTGCCCTCAACGCTGCCGACCTGTTGATTGGCG			778
Db 877	SSSKSSASSSVSSGSGSSGSSNSSASKSSSVSSGSGSGSGSVS			924

RESULT 3	AV390505	535 bp	mRNA	EST	21-JAN-2000
LOCUS	AV390505				
DEFINITION	AV390505	Chlamydomonas reinhardtii	C9	Chlamydomonas reinhardtii	
ACCESSION	AV390505	CDNA clone CM052B05_L		mRNA sequence.	
VERSION	AV390505.1	GI:5544721			
KEYWORDS	EST.				
SOURCE	Chlamydomonas reinhardtii.				
ORGANISM	Chlamydomonas reinhardtii				
	Eukaryota; Viridiplantae; Chlorophyta.				
	Chlamydomonadales; Chlamydomonas.				
REFERENCE	1 (bases 1 to 535)				

AUTHORS Asamizu, E., Nakamura, Y., Sato, S., Fukuzawa, H. and Tabata, S.
TITLE A Large Scale Structural Analysis of cDNAs in a Unicellular Alga, Chlamydomonas reinhardtii. I. Generation of 3433 Non-Redundant Expressed Sequence Tags
JOURNAL DNA Res. 6, 369-373 (1999)
COMMENT On Dec 20, 1995 this sequence version replaced g1:1135919.

FEATURES

Contact: yasukazu Nakamura
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: ynakamu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/Plant>
Location/Qualifiers

```

source
1..335
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/clone="CM052b05_r"
/clone_1b="Chlamydomonas reinhardtii C9"
/dev_stage="photoautotrophic growth"
/note="Vector: pBluescript SK+; Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT
107 a 195 c 148 g 85 t
ORIGIN

```

Query Match	4.2%	Score 51.8	DB 74	Length 535
Best Local Similarity	51.5%	Pred. No. 0.62		
Matches 119	Conservative 0	Mismatches 112	Indels 0	Gaps
QY 571	GCGCAGCTCGTGTGATCGCGCGCGGACCGCGGCTACCAAGCAGGCCATCGCCAC	630		
Db 221	GCTCAGCAGGAGGCGCATCGAGACTGCTCTCTTCACAGTACAGAGGCTTACCGGGGTG	280		
QY 631	GGCATGGGCGCGACCGCTTACGGTTCTTGACATCAACATCGAACATTTGGCAACTGCAC	650		
Db 281	CGCCTGGGCTGGACCGCTGGTGCCCAAGGCCCTCAAGTACCCAAAGCGGACGGGTGCAC	340		
QY 691	GCCGAGTTCTGGCGCGCGGACACACGCTGCTACATCGATCGGCTTACGAGCGAGGGGTGCC	750		
Db 341	CGCGACTGGAACCGCGCTCATGACCTCGTTCACAGCGGCGCTTCAACATCTGTGCAAGCG	400		
QY 751	GTCAAAGCTGCCGACCTGTGATTGGGCGCGCTCTGTGTGCGAGCGCCAG	801		
Db 401	GCGCGCTGGCGCTCCCTGCAAGCGGAGGGGCTCAAGATGAAGCCGCTAG	451		
RESULT 4				
CNS0052P/c	844 bp	UNA	GSS	03-JUN-1999
LOCUS	Drosophila melanogaster genome survey sequence TET3 end of BAC #			
DEFINITION	BACR1P16 of RFL1-98 library from Drosophila melanogaster (fruit			
	fly) genomic survey sequence.			
ACCESSION	AL036652			
VERSION	AL036652.1	GI:4932342		
KEYWORDS	GSS.			
SOURCE	fruit fly.			
ORGANISM	Drosophila melanogaster			
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;			
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
REFERENCE	1 (bases 1 to 844)			
AUTHORS	Genoscope.			
TITLE	Direct Submission			
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage			
	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)			
COMMENT	- Web : www.genoscope.cns.fr)			
	Determination of this BAC-end sequence was carried out as part of			
	collaboration with the Berkeley Drosophila Genome Project (BDGP).			
	The BDGP is constructing a physical map of the Drosophila			
	melanogaster genome using these BACs. For further information			
	please see http://www.fruitfly.org The BDGP Drosophila			
	melanogaster BAC library was prepared by Kazuhiro Osoegawa and			
	Aaron Mammosser in Pieter de Jong's laboratory in the Department of			


```

FEATURES
  Source
    Location/Qualifiers
      1.846
      /organism="Drosophila melanogaster"
      /plasmid="pBelobAC11"
      /db_xref="taxon:7227"
      /clone-lib="DrosBAC"
      /clone="BACN04N13"
      /note="end : 17"

BASE COUNT
  300 a 47 c 40 g 175 t 284 others

ORIGIN

Query Match
  3.7% Score 46.6; DB 82; Length 846;
  Best Local Similarity 13.0%; Pred. No. 6.7;
  Matches 33; Conservative 127; Mismatches 93; Indels 0; Gaps 0;

QY 369 TCGTTCACACGATGCGTTCGATTCGCGACACGATCAATGCTTACGAGACGCTCA 428
DB 842 TCTCTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSST 783
QY 429 GACCCGCGAGCGGCGACACCTGCTTCCCGGATGACCGAAGTCGCGGTCGACTCGC 488
DB 782 SSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 723
QY 489 CGCCGAGTGGCGCTTACCACTGATGGAACCCAGGCGCGCGTCTGATGGG 548
DB 722 SSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSST 663
QY 549 CGGGGTGCGCGGCGTCAACCGCGACGCTGATGATGCGCGCGCGCGCTCA 608
DB 662 SSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSST 603
QY 609 CAACGACGCGCGC 621
DB 602 CGCCGCGCGCGCGC 590

RESULT 7
CNS0181E/c 1101 bp DNA GSS 26-JUL-1999
LOCUS Drosophila melanogaster genome survey sequence SP6 end of BAC
DEFINITION BAC37H05 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL108764.1 GI:5629068
VERSION AL108764.1 GI:5629068
KEYWORDS fruit fly.
SOURCE Drosophila melanogaster
ORGANISM Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

COMMENT
  - Web : www.genoscope.cns.fr)
  Determination of this BAC-end sequence was carried out as part of a
  collaboration with the European Drosophila Genome Project (BDGP) -
  http://www.edgp.ebi.ac.uk/. This Drosophila melanogaster BAC
  library (Dros BAC) was made by Alain Billaud at CEPH (Centre
  d'Etude du Polymorphisme Humain) with funding provided by a MRC
  project grant. The DNA was prepared from embryos by Alain Bucheton
  and Genevieve Payan. It has been constructed in the vector
  pBelobAC11.

FEATURES
  Source
    Location/Qualifiers
      1..1101
      /organism="Drosophila melanogaster"
      /plasmid="pBelobAC11"
      /db_xref="taxon:7227"
      /clone-lib="DrosBAC"
      /clone="BACN37H05"

BASE COUNT
  257 a 170 c 162 g 96 t 250 others

```

```

BASE COUNT
  61 a 215 c 301 g 235 t 289 others

ORIGIN

Query Match
  3.7% Score 46.6; DB 83; Length 1101;
  Best Local Similarity 40.2%; Pred. No. 6.9;
  Matches 101; Conservative 40; Mismatches 109; Indels 1; Gaps 1;

QY 395 CCGGACACGATGCGCTTACGAGACGCGTCCAGACCGCGGCGGCGACACCGCTCC 454
DB 308 CCCGACVYMYVGMMAASCCCAVASCCSCCCSCSCGCGCCGCGCCGCGCGC 249
QY 455 TTGCCCCGATGCGGAGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 514
DB 248 AAGMGSMGSMGAGGACAGG GAGAGACCGCGGCGGCGGCGGCGGCGGCGGCGG 189
QY 515 TCGGACACCGAGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 574
DB 188 CGCGACCGCCSSAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 130
QY 575 AGCTGTGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCA 634
DB 129 ACGMAAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCA 70
QY 635 TGGCGCGGAC 645
DB 69 AGCGCGCAAC 59

RESULT 8
CNS006XK 935 bp DNA GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC #
DEFINITION BAC14N09 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL066051.1 GI:4945019
VERSION AL066051.1 GI:4945019
KEYWORDS fruit fly.
SOURCE Drosophila melanogaster
ORGANISM Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 935)
AUTHORS Genoscope.
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

COMMENT
  - Web : www.genoscope.cns.fr)
  Determination of this BAC-end sequence was carried out as part of a
  collaboration with the Berkeley Drosophila Genome Project (BDGP).
  The BDGP is constructing a physical map of the Drosophila
  melanogaster genome using these BACs. For further information
  please see http://www.fruitfly.org The BDGP Drosophila
  melanogaster BAC library was prepared by Kazuo Osoegawa and
  Aaron Mammosser in Pieter de Jong's laboratory in the Department of
  Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
  NY. The library is named RPCI-98 and was constructed by partial
  EcoRI digestion of Drosophila DNA provided by the BDGP from the
  isogenic strain y2; cn bw sp. The same strain used for the BDGP's
  pl and Est libraries. A more detailed description of the library
  and how to order individual BAC clones, the entire library, or
  filters for hybridization from the BACPAC Resource Center can be
  found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
  Source
    Location/Qualifiers
      1..935
      /organism="Drosophila melanogaster"
      /db_xref="taxon:7227"
      /clone-lib="RPCI-98"
      /clone="BACR14N09"
      /note="end : 17"

BASE COUNT
  257 a 170 c 162 g 96 t 250 others

```


DQ 335 TCCTTGCACGTTCTTCGATTTGGCCCCGTCAGTGCTGTGCACCAGTAGGTGTGGATT 394
 : : | | | : : : | | | : : : | | |
 DB 191 YCCTTTGCCCTTYSGGCGYSSGCCYCICGTCSSYBYHYTTGCCSAGCCTCNVSCCGTT 250

CCGGCACCAGTAAATGCTACGAGACCGTCACAGACGGCCGACGGACTACCCCTGC

Db 251 yccctcvcyyvyygytccccccscggccgcvtccccggccccccctcccccscuccccccccccccccccccc sit

[illegible]

515 TCGGAACCCAGGCGGGCCGCGGTGCTGATGGCGGGGTGCCCGGCGTGAACCGGCCG 574

Db 371 CSCGNGCNGSSGGCCSSSSSGSGSSSSGGGGGGCGGSCNGSCGCGGGGNGNGGSG 430

575 ACGTCGTGTGATCGG 590

Db 431 NGGNGSSGSKTKG 445

Result 13

CNS017SY	LOCUS	CNS017SY	1101 bp	DNA	GSS	26-JUL-1999

DEFINITION *Drosophila melanogaster* genome survey sequence Sp6 end of BAC BACN37L08 of Drosophila library from *Drosophila melanogaster* (fruit fly)

Accession	fly), genomic survey sequence.	
AL108460		
U108460	CT-538764	9

KEYWORDS	GSS:
VERSION	AL108400.1
SOURCE	fruit fly
	GI:5026704

ORGANISM Drosophila melanogaster
Eukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta:

Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.

TITLE Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequen-
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequen-
nd 101 01006 envy codex - FRANCE (E-mail : secref@genoscope.C

COMMENT
BP 131 31008 ENVI CECHA FRANCE (L MAIL : degrange@genoscope.fr)
- Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as par

collaboration with the European Drosophila Genome Project (EDGP) <http://www.edgp.ebi.ac.uk>. This *Drosophila melanogaster* BAC

library (Dros-BAC) was made by Alain Billard at CERH (Centre d'Etude du Polymorphisme Humain) with funding provided by a M

project grant. The DNA was prepared from embryos by Alain Buc and Genevieve Payan. It has been constructed in the vector

FEATURES

Source	Location/Qualifiers
1	1101

```
source
1: read
/organism="Drosophila melanogaster"
/plasmid="pBelioBAC11"
```

```

/2_
/db_xref="taxon:7227"
/clone_lib="DrosBAC"

```

```
/clone="BACN37L08"  
/note="end : Sp6"
```

BASE COUNT	254 a	176 c	160 g	152 f	359 Others
ORIGIN					

Query Match 3.6% Score 45.2 DB 83 Length 1101

Best Local Similarity 12.6%; Pred. No. 13;
Matches 40; Conservative 157; Mismatches 121; Indels 0; Gaps

426 CCAGACGGCGAGCGGCGCACTACCCCTGCTTGGCCCGATGAGCGGAAGTGGCCGCTGACT 48

Db 755 CGAGAGGGGSGRGGGGGCCASSGVCACACSSASACSSASAMCGVSSGGSSCASCGGCCGVSS 81

486 CGCCCCCAGGTTGGCGCTTACCACCTGATGCGAACCAGGGGGCCGGGTGTCTGAT 54

Db 815 CSAVASASSSTYMSKVASAVASCASAVASGMSAGAVSSSCRSVASVSAASVSSSVSSS 87

•

QY 546 GGGGGGGTCCCGCGCGTGAACCGCGCGCGTGTGATCGCGCGCGCGCGCG 605
 Db 875 SSSSVSSAAASASSSSSSAAVAAAASVAVSSVSSSSSSSSSSSSSSSSSS 934
 QY 606 CTACACCGACCGCGCGTGAACCGCGCGCGTGTGATCGCGCGCGCGCG 665
 Db 935 SVAAASASSSSSSSSSVSSSVSSSVSSSVSSSVSSSVSSSVSSSVSS 994
 QY 666 CATGCAACACTTGGCGCACTGACGCGCGTGTGATCGCGCGCGCGTGTG 725
 Db 995 AAAAAAASASSSSAAVAVSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 1054
 QY 726 ATCGCGCTACGAGCTCGA 743
 Db 1055 VSAVASASASVSSVA 1072

RESULT 13
 A1712257 315 bp mRNA EST 02-FEB-2000
 LOCUS 605069D09.x1 605 - Endosperm cDNA library from Schmidt lab Zea mays
 DEFINITION cDNA, mRNA sequence.

ACCESSION A1712257
 VERSION A1712257.1 GI:5006195
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 eudicotyledons; Euphyllophytes; Spermatophyta; Magnoliophyta; Liliops; Poales;
 Poaceae; Zea.
 1 (bases 1 to 315)
 Mailbot V.
 Mailze ESTs from various cDNA libraries sequenced at Stanford
 University.
 Unpublished (1999)
 On Jun 22, 1998 this sequence version replaced gi:3246668.
 CONTACT: Mailbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: mailbot@stanford.edu
 Plate: 605069 row: D column: 09.
 Location/Qualifiers

FEATURES
 source 1. 315
 /organism="Zea mays"
 /cultivar="Ohio43"
 /db_xref="taxon:4577"
 /clone_lib="605 - Endosperm cDNA library from Schmidt lab"
 /tissue_type="nucellar, embryo, and endosperm"
 /dev_stage="10-14 days post-pollination"
 /lab_host="DH5(alpha)"
 /note="Organ: Kernel; Vector: PAD-GAL4-27; Site: 1; ECORI;
 Site: XhoI; Kernel endosperm cDNA library from Schmidt
 lab"

BASE COUNT 61 a 105 c 106 g 42 t 1 others
 ORIGIN

Query Match 3.6%; Score 45; DB 51; Length 315;
 Best Local Similarity 50.7%; Pred. No. 12;
 Matches 108; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 397 GGCACGACGATGCTGCTACGACGCGTCAAGCGCGCGCGCGCGCGCTT 456
 Db 58 GGCATGCGCTCTCCATCAAAAGACAGCTGGTGGCGAGCAACTTAAACACCC 117
 QY 457 GCGCCGATGAGCAAGTGGCGGCTGCACTGCGCGCGCGCGCGCTTACCACTGATG 516
 Db 118 GCGCTTAGCAGCAAGCGCGGCTCAAGCGGATAGCGCGCGCGCGCGCGCGCGCG 177

QY 517 CGAACCAAGGGGGCGCGCGTGTGATGCGCGCGCGCGCGCGCGCGCGAC 576
 Db 178 CGTGGCGGACGCGCGCGCGCGCGCGTGTGATGCGCGCGCGCGCGCGCG 237
 QY 577 GTGTGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 609
 Db 238 GTGGCGGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 270

RESULT 14
 CNS010R1 846 bp DNA GSS 26-JUL-1999
 LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC
 DEFINITION BACND04N13 of Drosophila library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.

ACCESSION AL099337
 VERSION AL099337.1 GI:5610948
 KEYWORDS GSS.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 846)
 Genoscope.
 Direct Submission
 Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage;
 BP 191 91006 EVRY cedex - FRANCE (E-mail: sequef@genoscope.cns.fr
 - Web: www.genoscope.cns.fr)

COMMENT
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the European Drosophila Genome Project (EDGP) -
 http://www.edgp.edi.ac.uk - This Drosophila melanogaster BAC
 library (Dros BAC) was made by Alain Billard at CEPH (Centre
 d'Etude du Polymorphisme Humain) with funding provided by a MRC
 project grant. The DNA was prepared from embryos by Alain Bucheton
 and Genevieve Payan. It has been constructed in the vector
 pBeloBac11.

FEATURES
 source 1. 846
 Location/Qualifiers

/organism="Drosophila melanogaster"
 /plasmid="pBeloBac11"
 /db_xref="taxon:7227"
 /clone_lib="DrosBAC"
 /clone="BACND04N13"
 /note="end: T7"

BASE COUNT 300 a 47 c 40 g 175 t 284 others
 ORIGIN

Query Match 3.6%; Score 45; DB 82; Length 846;
 Best Local Similarity 16.2%; Pred. No. 14;
 Matches 38; Conservative 109; Mismatches 88; Indels 0; Gaps 0;

QY 512 TGATGCAACGAGGGGGCGCGGTGTGATGCGCGCGCGCGCGCGCGCGCG 571
 Db 574 KGACGSGAARSAGAGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 693

QY 572 CGAGCTGCTGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 651
 Db 634 SSGSSSSAAASAGASASVSSASASVSSSSSSASASASVSSSVAVASASSS 693

QY 632 GCATGGCGCGACCGCTTACGATCAACATGCAACATGCGCGCGCGCGCG 691
 Db 694 SSVSVSSSSAAVAVASASASVSSSVSSSVSSSVSSSVSSSVSSSVSS 753

QY 692 CCGATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 746
 Db 754 SSASASSSSAAASASSSSAAASASSSSAAASASSSSAAASASSSSSS 808

RESULT 15
 CNS0072Q 932 bp DNA GSS 03-JUN-1999
 LOCUS

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 14:50:58 ; Search time 8627.09 Second

(Without alignments)
-139.259 Million cell updates/sec

Title: US-09-362-485-3
Perfect score: 1235
Sequence: 1 ATCTTCAGATTAATCGAAC.....ANGAAGCGATGATCGTGCNC /235

Scoring table: IDENTITY_NTC
Gapop 10.0 , Gapext 1.0

Searched: 882769 segs, -486395729 residues

Total number of hits satisfying chosen parameters: 1765534

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_da1:*
2: gb_da2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl3:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vl:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vl:*
32: gb_hcg1:*
33: gb_hcg2:*
34: gb_hcg3:*
35: gb_hcg4:*
36: gb_hcg5:*
37: gb_hcg6:*
38: em_hum3:*
39: em_hum4:*
40: gb_pr4:*
41: gb_hcg3:*
42: gb_hcg4:*
43: gb_hcg5:*
44: gb_hcg6:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1232	99.8	1235	5	A87606 Sequence 3
2	1232	99.8	1235	5	A89746 Sequence 3
3	1232	99.8	1236	5	A87613 Sequence 10
4	1232	99.8	1236	5	A89753 Sequence 10
5	1230.4	99.6	1237	5	A87607 Sequence 4
6	1230.4	99.6	1237	5	A89747 Sequence 4
7	1230.4	99.6	56414	1	MTW0002
8	1224.4	99.1	1228	5	A87608 Sequence 5
9	1224.4	99.1	1228	5	A89748 Sequence 5
10	1219	98.7	1235	5	A87609 Sequence 6
11	1219	98.7	1235	5	A89749 Sequence 6
12	1218.4	98.7	1235	5	A87611 Sequence 8
13	1218.4	98.7	1235	5	A89751 Sequence 8
14	1216	98.5	1245	5	A87605 Sequence 2
15	1216	98.5	1245	5	A89745 Sequence 2
16	1216	98.5	1260	5	A87604 Sequence 1
17	1216	98.5	1260	5	A89744 Sequence 1
18	1216	98.5	2584	1	MTALADH
19	1213.4	98.3	1229	5	A87610 Sequence 7
20	1213.4	98.3	1229	5	A89750 Sequence 7
21	1207	97.7	1209	5	A87612 Sequence 9
22	1201	97.2	1208	5	A89752 Sequence 9
23	1194	96.7	1194	2	MTU92472
24	698	56.5	720	5	A89759 Sequence 16
25	538.4	43.6	40745	1	SC151
26	467.6	37.9	682	5	A89758 Sequence 15
27	467.6	37.9	682	5	A89762 Sequence 19
28	377	30.5	8046	2	AF049107
29	351.2	28.4	11514	2	AE002028
30	351.2	26.8	2780	1	RLE238118
31	330.8	26.8	1585	2	AF070716
32	312.8	25.3	1736	1	BACALDHC
33	293	23.7	1477	2	AF070715
34	288.4	23.4	2113	2	D37807
35	277.2	22.4	1125	5	E01847
36	258	20.9	13740	1	D90800
37	243.2	19.7	1375	1	AB013821
38	243	19.7	1776	1	BACALDH
39	235.6	19.1	1709	1	BACALD
40	235.6	19.1	12316	1	BS282015
41	235.6	19.1	217420	1	BSUB0017
42	218.2	17.7	1714	2	AF070714
43	107.2	8.7	10652	2	AE001565
44	99.2	8.0	10512	2	AE000639
45	74.2	6.0	3160	1	RR05294

ALIGNMENTS

```

RESULT 1
AB7606 1235 bp DNA PAT 22-JAN-2000
LOCUS A87606
DEFINITION Sequence 3 from Patent WO9836089.
ACCESSION A87606
VERSION A87606.1 GI:6736246
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1235)
AUTHORS Flohe, L. and Singh, M.
TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL Patent: WO 9836089-A 20-AUG-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
source 1..1235
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 236 a 394 c 382 g 220 t 3 others
ORIGIN
Query Match 99.8%; Score 1232; DB 5; Length 1235;
Best Local Similarity 100.0%; Pred. No. 3.3e-169;
Matches 1235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATCTTGAGATTATGCAACTTTCTTCACTGAGCTAGCATGATGAGAGGGTATC 60
DB 1 ATCTTGAGATTATGCAACTTTCTTCACTGAGCTAGCATGATGAGAGGGTATC 60
QY 61 ATGCGCTGCTGATTCGACGAGACAAAAACAAGCAATTCGGGTGGCCATACCCG 120
DB 61 ATGCGCTGCTGATTCGACGAGACAAAAACAAGCAATTCGGGTGGCCATACCCG 120
QY 121 GCGGCGTGGGGAAGCAAGCCGCTGATGAGTGTCTATCCAGCAGATGCCGA 180
DB 121 GCGGCGTGGGGAAGCAAGCCGCTGATGAGTGTCTATCCAGCAGATGCCGA 180
QY 181 GAGGCTCGGCTATACCGAGCGGATTTCAAGCGCGGCGGCACTGTTGGGACC 240
DB 181 GAGGCTCGGCTATACCGAGCGGATTTCAAGCGCGGCGGCACTGTTGGGACC 240
QY 241 GCGGCGAGGTGGGGCGAGCTGATTTATGCTCAAGTCAAGCAAGCAAGCGGCG 300
DB 241 GCGGCGAGGTGGGGCGAGCTGATTTATGCTCAAGTCAAGCAAGCAAGCGGCG 300
QY 301 GAATACGCGCGCTGCGACAGGCGAGATCTTGTACGTTCTTCATTTGGCGCTCA 360
DB 301 GAATACGCGCGCTGCGACAGGCGAGATCTTGTACGTTCTTCATTTGGCGCTCA 360
QY 361 CGTGTGACCGATGCTGTTGATTCGGGACACGATGATGCTGATGCTGATGCTG 420
DB 361 CGTGTGACCGATGCTGTTGATTCGGGACACGATGATGCTGATGCTGATGCTG 420
QY 421 CAGACCGCGAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB 421 CAGACCGCGAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 481 GCGGCGAGGTGGGCTTACCACTGATGCGAAGCAAGGCGGCGGCTGCTGCTGATG 540
DB 481 GCGGCGAGGTGGGCTTACCACTGATGCGAAGCAAGGCGGCGGCTGCTGCTGATG 540
QY 541 GCGGCGAGGTGGGCTTACCACTGATGCGAAGCAAGGCGGCGGCTGCTGCTGATG 600
DB 541 GCGGCGAGGTGGGCTTACCACTGATGCGAAGCAAGGCGGCGGCTGCTGCTGATG 600
QY 601 TACAAGCGACCGCGCATGCGCAAGCGGCGGCGGCGGCGGCTTACGCTTACGATCAAC 660
DB 601 TACAAGCGACCGCGCATGCGCAAGCGGCGGCGGCGGCGGCTTACGCTTACGATCAAC 660
QY 661 ATGACCAAACTTGGGCACTGACGCGGAGTGTGGGCGGATTCACACCTGCTACTCA 720

```

```

DB 661 ATGACCAAACTTGGGCACTGACGCGGAGTTCGCGGCGGATCCACTGCTACTCA 720
QY 721 TCGGCGTACGAGCTCGAGGCGTCCGTCACAAAGTCCGACGCTGATGAGGCGGCTGCT 780
DB 721 TCGGCGTACGAGCTCGAGGCGTCCGTCACAAAGTCCGACGCTGATGAGGCGGCTGCT 780
QY 781 GTGCGAGGCGGCGCAAGGCGGCAAAATAGTCTGCAATTCCTGCTGCGCATATGAACCA 840
DB 781 GTGCGAGGCGGCGCAAGGCGGCAAAATAGTCTGCAATTCCTGCTGCGCATATGAACCA 840
QY 841 GGTGCGTACGAGCTCGATATGATGATGATGATGATGATGATGATGATGATGATGATG 900
DB 841 GGTGCGTACGAGCTCGATATGATGATGATGATGATGATGATGATGATGATGATGATG 900
QY 901 ACCACCTAGACACACCGGCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
DB 901 ACCACCTAGACACACCGGCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
QY 961 ATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 961 ATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 1021 GTGCTGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
DB 1021 GTGCTGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
QY 1081 AAAGCTTTTGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
DB 1081 AAAGCTTTTGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
QY 1141 GTGCGCTTACCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
DB 1141 GTGCGCTTACCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
QY 1201 CACACNTGCGGAGTAAGC. GCGATGATGTCGNC 1235
DB 1201 CACACNTGCGGAGTAAGCAGATGATGTCGNC 1235
RESULT 2
AB9746 1235 bp DNA PAT 22-JAN-2000
LOCUS A89746
DEFINITION Sequence 3 from Patent WO9832862.
ACCESSION A89746
VERSION A89746.1 GI:6738280
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1235)
AUTHORS Flohe, L. and Singh, M.
TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
JOURNAL Patent: WO 9832862-A 30-JUL-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
source 1..1235
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 236 a 394 c 382 g 220 t 3 others
ORIGIN
Query Match 99.8%; Score 1232; DB 5; Length 1235;
Best Local Similarity 100.0%; Pred. No. 3.3e-169;
Matches 1235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATCTTGAGATTATGCAACTTTCTTCACTGAGCTAGCATGATGAGAGGGTATC 60
DB 1 ATCTTGAGATTATGCAACTTTCTTCACTGAGCTAGCATGATGAGAGGGTATC 60
QY 61 ATGCGCTGCTGATTCGACGAGACAAAAACAAGCAATTCGGGTGGCCATACCCG 120

```

Db 61 ATGCGGCTGCGTATTCGACCGAGACCAAAACAGCAATTCGCGGTGGCCATCACCCG 120
QY 121 GCCGGCTGCGGAACTAACCCCTCGTGGCCATGAGTGTCTATCCAGCAGTGGCCGA 180
Db 121 GCCGGCTGCGGAACTAACCCCTCGTGGCCATGAGTGTCTATCCAGCAGTGGCCGA 180
QY 181 GAGGGCTGCGGCTATACCGAGCGGATTTCAAGCGGCGAGCGCGCAACTGCTGGCACC 240
Db 181 GAGGGCTGCGGCTATACCGAGCGGATTTCAAGCGGCGAGCGCGCAACTGCTGGCACC 240
QY 241 GCCGACGAGGTGTGGCGCGAGCGGTATTTATGCTCAAGGTCAAGAACCGATAGCGCG 300
Db 241 GCCGACGAGGTGTGGCGCGAGCGGTATTTATGCTCAAGGTCAAGAACCGATAGCGCG 300
QY 301 GAATAGCGCGCTGCGACACAGCGGAGATTTGTTACGTTCTTGCAATTTGGCCGCTCA 360
Db 301 GAATAGCGCGCTGCGACACAGCGGAGATTTGTTACGTTCTTGCAATTTGGCCGCTCA 360
QY 361 CGTGTGACCGAGTGGTGTGGATTTCGGGACACAGTCAATGGCTTACGAGACCGCTC 420
Db 361 CGTGTGACCGAGTGGTGTGGATTTCGGGACACAGTCAATGGCTTACGAGACCGCTC 420
QY 421 CAGACCGCGACGCGACCTACCTGCTTGCCTGATGA 3GAAATGCGCGGTGACCTC 480
Db 421 CAGACCGCGACGCGACCTACCTGCTTGCCTGATGA 3GAAATGCGCGGTGACCTC 480
QY 481 GCGCGCCAGTGTGGCGCTTACACCTGATGCGAACCCAAAGGGGGCGCGGTGTCTATG 540
Db 481 GCGCGCCAGTGTGGCGCTTACACCTGATGCGAACCCAAAGGGGGCGCGGTGTCTATG 540
QY 541 GGGGGGTGCGCGCGCTGCAACGCGGCGAGTGGTGTGATGGCGCGCGGACCGCGCG 600
Db 541 GGGGGGTGCGCGCGCTGCAACGCGGCGAGTGGTGTGATGGCGCGCGGACCGCGCG 600
QY 601 TACAACGACCGCGCATCGCCAAAGCGCATGGCGCGACCGCTTACGTTCTAGACATCAAC 660
Db 601 TACAACGACCGCGCATCGCCAAAGCGCATGGCGCGACCGCTTACGTTCTAGACATCAAC 660
QY 661 ATGACAACTTGGGCACTGACCGCGAGTTCGCGCGGAGTCCACACTCGTACTCA 720
Db 661 ATGACAACTTGGGCACTGACCGCGAGTTCGCGCGGAGTCCACACTCGTACTCA 720
QY 721 TCGGCGTACGAGTGGGCGGTGCGCAACGCTGCGAGTGGGCGGTGCTG 780
Db 721 TCGGCGTACGAGTGGGCGGTGCGCAACGCTGCGAGTGGGCGGTGCTG 780
QY 781 GTGCGACGCGCAAGGCAACCAATTAATGCTGCAATTCACCTGCGCGCATATGAACA 840
Db 781 GTGCGACGCGCAAGGCAACCAATTAATGCTGCAATTCACCTGCGCGCATATGAACA 840
QY 841 GTGCGAGTGTGATATGACCATGACAGCGCGGTGCTGCAAGGCTCACGACCG 900
Db 841 GTGCGAGTGTGATATGACCATGACAGCGCGGTGCTGCAAGGCTCACGACCG 900
QY 901 ACCACCTACGACCGAGGCTGCGGAGTGGCGAGACAGCGGCTTACTGCTGGGCAAC 960
Db 901 ACCACCTACGACCGAGGCTGCGGAGTGGCGAGACAGCGGCTTACTGCTGGGCAAC 960
QY 961 ATGCGCGCTGCGTGGAGAGACGTGACCTACGCGGTGACCAAGCGAGATGCGCTAT 1020
Db 961 ATGCGCGCTGCGTGGAGAGACGTGACCTACGCGGTGACCAAGCGAGATGCGCTAT 1020
QY 1021 GTGCTGAGTGTGGGAGCAATGAGTGGGCGGCGGTGCGGAGTGGGCGCTATGCG 1080
Db 1021 GTGCTGAGTGTGGGAGCAATGAGTGGGCGGCGGTGCGGAGTGGGCGCTATGCG 1080
QY 1081 AAAGGTCTTTCGACGCAAGAGGGGCTTACTGTCGAACGCGGTGCGACGACCTGGGG 1140
Db 1081 AAAGGTCTTTCGACGCAAGAGGGGCTTACTGTCGAACGCGGTGCGACGACCTGGGG 1140
QY 1141 GTGCGCTTACGAGCGCGCGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
Db 1141 GTGCGCTTACGAGCGCGCGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200

QY 1201 CACACNTCGGAGTAANGAGAGCATGATGTGNC 1235
Db 1201 CACACNTCGGAGTAANGAGAGCATGATGTGNC 1235
RESULT 3
A87613 1236 bp DNA PAT 22-JAN-2000
LOCUS Sequence 10 from Patent WO9836089.
DEFINITION
ACCESSION A87613
VERSION A87613.1 GI:5736253
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1236)
Flobe, L. and Singh, M.
AUTHORS
TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL Patent: WO 9836089-A 20-AUG-1998;
FLOBE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
source
1..1236
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 236 a 395 c 385 g 220 t
ORIGIN
Query Match 99.8%; Score 1232; DB 5; Length 1236;
Best local similarity 99.8%; Pred. No. 3.3e-169;
Matches 1232; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATCTTCAGATTATGAACCTTCTTACACTGAGAGCTATGAGAGGGTAAATC 60
Db 1 ATCTTCAGATTATGAACCTTCTTACACTGAGAGCTATGAGAGGGTAAATC 60
QY 61 ATGCGGCTGCGTATTCGCGACCGAGACCAAAACAGCAATTCGCGGTGGCCATCACCCG 120
Db 61 ATGCGGCTGCGTATTCGCGACCGAGACCAAAACAGCAATTCGCGGTGGCCATCACCCG 120
QY 121 GCCGCGTGGGAGTAACCCGCTGCGCATGAGTGGCTATCCAGGAGGTGGCGGA 180
Db 121 GCCGCGTGGGAGTAACCCGCTGCGCATGAGTGGCTATCCAGGAGGTGGCGGA 180
QY 181 GAGGGCTGCGCTATCACCGAATTCAGAGCGGCGAGCGGCACTGCTGGCACC 240
Db 181 GAGGGCTGCGCTATCACCGAATTCAGAGCGGCGAGCGGCACTGCTGGCACC 240
QY 241 GCCGACGAGTGTGGCGCGAGCTGATTTATGCTCAAGTCAAAAGCGATAGCGGG 300
Db 241 GCCGACGAGTGTGGCGCGAGCTGATTTATGCTCAAGTCAAAAGCGATAGCGGG 300
QY 301 GAATAGCGCGCTGCGACACAGCGGAGATTTGTTACGTTCTTGCAATTTGGCCGCTCA 360
Db 301 GAATAGCGCGCTGCGACACAGCGGAGATTTGTTACGTTCTTGCAATTTGGCCGCTCA 360
QY 361 CGTGTGACCGAGTGGTGTGGATTTCGGGACACAGTCAATGGCTTACGAGACCGCTC 420
Db 361 CGTGTGACCGAGTGGTGTGGATTTCGGGACACAGTCAATGGCTTACGAGACCGCTC 420
QY 421 CAGACCGCGAGCGGAGTACCGCTGCTGGCGGAGTGGGAGTGGGAGTGGGAGTGG 480
Db 421 CAGACCGCGAGCGGAGTACCGCTGCTGGCGGAGTGGGAGTGGGAGTGGGAGTGG 480
QY 481 GCGCGCGAGTGGCGCTTACACTGATGCGAACCAAGGGGGCGCGGTGTGCTGATG 540
Db 481 GCGCGCGAGTGGCGCTTACACTGATGCGAACCAAGGGGGCGCGGTGTGCTGATG 540
QY 541 GCGCGCGAGTGGCGCTTACACTGATGCGAACCAAGGGGGCGCGGTGTGCTGATG 600
Db 541 GCGCGCGAGTGGCGCTTACACTGATGCGAACCAAGGGGGCGCGGTGTGCTGATG 600

Qy	601	TTCACGCGACGCCGCATCGCCAAAGCGCATGGCGCGACCGCTACGGTCTAGACATCAAC	601
Db	601	TTCACGCGACGCCGCATCGCCAAAGCGCATGGCGCGACCGCTACGGTCTAGACATCAAC	666
Qy	661	ATCGACAACCTTCGGCAATCTCGACCCCGAGTTCTCGCGCGCGATCCACACTCGCTACTCA	721
Db	661	ATCGACAACCTTCGGCAATCTCGACCCCGAGTTCTCGCGCGCGATCCACACTCGCTACTCA	722
Qy	721	TGGGCTACGAGCTCGAGGGTGCCTCAAAAGTGGCCGACCTTGATTTGGGGCGCTCTG	781
Db	721	TGGGCTACGAGCTCGAGGGTGCCTCAAAAGTGGCCGACCTTGATTTGGGGCGCTCTG	782
Qy	781	GTGGCAGGGCCCAAGGACCCAAATTAAGTCTGGAATTAAGTCTGGCCATATGAACA	841
Db	781	GTGGCAGGGCCCAAGGACCCAAATTAAGTCTGGAATTAAGTCTGGCCATATGAACA	842
Qy	841	GGTGGCGTACTGTGTGATATAGCAATCGACACAGGGCGGCTGTTTGAAGGCTCACGACCG	901
Db	841	GGTGGCGTACTGTGTGATATAGCAATCGACACAGGGCGGCTGTTTGAAGGCTCACGACCG	902
Qy	901	ACCACCTAGACACACCCGAGCTTGCGCGCGACGACAGCGCTTTTACTGGCTGGCGAAC	961
Db	901	ACCACCTAGACACACCCGAGCTTGCGCGCGACGACAGCGCTTTTACTGGCTGGCGAAC	962
Qy	961	ATGCCCGCTCGGTGCGGAAAGACGTGACACTACGGCTGACCAAGCGACGATGCCGTAT	1011
Db	961	ATGCCCGCTCGGTGCGGAAAGACGTGACACTACGGCTGACCAAGCGACGATGCCGTAT	1012
Qy	1021	GTGGCTGACACTTGGCGACCATGGCTGGGGGGGGGGGGTGGCGGCTGCAATCGGCATACCC	1081
Db	1021	GTGGCTGACACTTGGCGACCATGGCTGGGGGGGGGGGGTGGCGGCTGCAATCGGCATACCC	1082
Qy	1081	AAAGGCTTTTCGACGACGACGAAGGGCGTTACTGTCCGAACGGGGTGGCCACCGACTGGGG	1141
Db	1081	AAAGGCTTTTCGACGACGACGAAGGGCGTTACTGTCCGAACGGGGTGGCCACCGACTGGGG	1142
Qy	1141	GTGGCTTACCGGAGCGCCCGACGGTGTGGCTGACTGCGCGCGTTCGGCCGCTTACGCCGAG	1201
Db	1141	GTGGCTTACCGGAGCGCCCGACGGTGTGGCTGACTGCGCGCGTTCGGCCGCTTACGCCGAG	1202
Qy	1201	CACACNTCGGAGTAGAAGGAACGATGATGTCNC	1235
Db	1201	CACACNTCGGAGTAGAAGGAACGATGATGTCNC	1236
RESULT	4		
LOCUS	A89753	1236 bp	DNA
DEFINITION	Sequence 10 from Patent WO9832862.		
ACCESSION	A89753		
VERSION	A89753.1	GI:6738287	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
BASE COUNT	236 a	395 c	385 g
ORIGIN			
Query Match	99.8%	Score 1232	DB 5: Length 1236;
Best Local Similarity	99.8%	Pred. No.3.3e-169;	
Matches 1232;	Conservative	0; Mismatches	3; Indels
0; Gaps			
1	ATCTTGCAGATTAAATCGAAGCTTCTTCACTGAAGCGTACAGTATCGAAGGGTAAATC	60	

D	b	1	ATCTTGACGATTAAATCGAACCTTTCTTCCACTGAACCGTACAGTATCGAGAGGGTAAATC	60
O	y	61	ATGGCGCTGCGTATATCCGACCGAGACCAAAAACACGAATTCGGGTGGCATACCCG	120
D	b	61	ATGGCGCTGGTATATCCGACCGAGACCAAAAACACGAATTCGGGTGGCATACCCG	120
O	y	121	GCCGGCGTCGCGGAACCTAACCCGTCGTGGCATAGAGTCTATCCAGGAGTCCGGA	180
D	b	121	GCCGGCGTCGCGGAACCTAACCCGTCGTGGCATAGAGTCTATCCAGGAGTCCGGA	180
O	y	181	GAGGCGTCGCGTATACCGACCGCGGATTTCAAGCGCGAGCGCACTGGTGGCAC	240
D	b	181	GAGGCGTCGCGTATACCGACCGCGGATTTCAAGCGCGAGCGCACTGGTGGCAC	240
O	y	241	GCCGACCAAGTGTGGGCCGACGCTGATTTATGTCTCAAGTCAAAAGAACCGATAGCGG	300
D	b	241	GCCGACCAAGTGTGGGCCGACGCTGATTTATGTCTCAAGTCAAAAGAACCGATAGCGG	300
O	y	301	GAATAGGCGCCGCTGGGACACCGGGACAGATCTGTAGCTCTGTGATTTGGCGCGTCA	360
D	b	301	GAATAGGCGCCGCTGGGACACCGGGACAGATCTGTAGCTCTGTGATTTGGCGCGTCA	360
O	y	361	CGTGCTTGACACCGATGCGGTTGTGGATTCGCGACCAAGTCATATGCTTACAGACCGT	420
D	b	361	CGTGCTTGACACCGATGCGGTTGTGGATTCGCGACCAAGTCATATGCTTACAGACCGT	420
O	y	421	CAGACCGCGGACGGCGACACACCCGCTGTGGCCGATAGGAAGTGGCGGTGCAATC	480
D	b	421	CAGACCGCGGACGGCGACACACCCGCTGTGGCCGATAGGAAGTGGCGGTGCAATC	480
O	y	481	GCCGCGCAGGTTGGCGCTTACCACTGATGTCGAACCAAGAGGGGGCCGCGTGTGATG	540
D	b	481	GCCGCGCAGGTTGGCGCTTACCACTGATGTCGAACCAAGAGGGGGCCGCGTGTGATG	540
O	y	541	GCGGGGGTGGCGCGCTGTCGAACCGCGGACAGTGTGTGATGGCGCGGACCGCGCG	600
D	b	541	GCGGGGGTGGCGCGCTGTCGAACCGCGGACAGTGTGTGATGGCGCGGACCGCGCG	600
O	y	601	TACAACGACGCCCATCTGCGCAACGGCATGGCGGCACGTTACGTTCTACATCAAC	660
D	b	601	TACAACGACGCCCATCTGCGCAACGGCATGGCGGCACGTTACGTTCTACATCAAC	660
O	y	661	ATCGACAACCTTGGGCAACTGCAACGCCGAGTTCTGCGCGGATCCACACTGCTACTCA	720
D	b	661	ATCGACAACCTTGGGCAACTGCAACGCCGAGTTCTGCGCGGATCCACACTGCTACTCA	720
O	y	721	TCGGGCTACGAGCTCGAGGGGTCGTCGAACGTCGCCAGCTGGTATATGGGGCGTCTG	780
D	b	721	TCGGGCTACGAGCTCGAGGGGTCGTCGAACGTCGCCAGCTGGTATATGGGGCGTCTG	780
O	y	781	GTCGACGCGCCAAAGGCAACCAATTAGTCTCAATCTACTGTGCGCATATGAACCA	840
D	b	781	GTCGACGCGCCAAAGGCAACCAATTAGTCTCAATCTACTGTGCGCATATGAACCA	840
O	y	841	GSTGCGTACTGTGTGATATAGCCATCGACACAGGGCGCTGTTGGAAGCTCACAGCG	900
D	b	841	GSTGCGTACTGTGTGATATAGCCATCGACACAGGGCGCTGTTGGAAGCTCACAGCG	900
O	y	901	ACCACTACGACACCGCGAGCTTGGCGTGCAGACAGCCTGTTTACTGGTGGCGAC	960
D	b	901	ACCACTACGACACCGCGAGCTTGGCGTGCAGACAGCCTGTTTACTGGTGGCGAC	960
O	y	961	ATGCGCGCTCGTGTGCGGAAGAGCTGACCTACGCGCTGACCAAGCGCAGATGCCGAT	1020
D	b	961	ATGCGCGCTCGTGTGCGGAAGAGCTGACCTACGCGCTGACCAAGCGCAGATGCCGAT	1020
O	y	1021	GTCGTCAGAGCTTGGCGACCAATGGCTGGCGGGCGGCTGGCGGTCAATTCGGCATAC	1080
D	b	1021	GTCGTCAGAGCTTGGCGACCAATGGCTGGCGGGCGGCTGGCGGTCAATTCGGCATAC	1080
O	y	1081	AAAGGCTTTTCGAGCAGCAAGAGGGGGTTCCTCTCGAAGGGGTGACCAACGACTTGGG	1140

```

Db 1081 AAGGCTTTTCGACGACGAAGGGCGTTACTGTCGGAAGGGGTGGCCACCGACCTGGGG 1140
QY 1141 GTGCGCTTACCGAGCCGCGAGCGTCTGGCTGACTCTCGCGCGTCTGTTACCGGAG 1200
Db 1141 GTGCGCTTACCGAGCCGCGAGCGTCTGGCTGACTCTCGCGCGTCTGTTACCGGAG 1200
QY 1201 CACACNTCGGGAGTAANGAAGCGGATGTCGNC 1235
Db 1201 CACACGTCGGGAGTAAGGAAGCGATGTCGCGC 1235

RESULT 5
AB7607 1237 bp DNA PAT 22-JAN-2000
LOCUS Sequence 4 from Patent WO9836089.
DEFINITION AB7607
ACCESSION AB7607.1 GI:6736247
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1237)
AUTHORS Flohe L. and Singh M.
TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL Patent: WO 9836089-A 20-AUG-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
source 1..1237
Location/Qualifiers
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 236 a 394 c 386 g 221 t
ORIGIN

Query Match 99.68; Score 1230.4; DB 5; Length 1237;
Best Local Similarity 99.78; Pired. No. 5; Ge-169;
Matches 1231; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATCTTGAGATTATCGAAGTTCTTCTCACACTGAGAGTACGATGAGAGGGTAAATC 60
Db 1 ATCTTGAGATTATCGAAGTTCTTCTCACACTGAGAGTACGATGAGAGGGTAAATC 60
QY 61 ATGCGCGTCGCTATTCGCGACGAGAACCAAAAACGAATTCGGGTGCCATACCCCG 120
Db 61 ATGCGCGTCGCTATTCGCGACGAGAACCAAAAACGAATTCGGGTGCCATACCCCG 120
QY 121 GCGGCGCTCGCGGAACAAACCCGCTGCGGCAATGAGTCTCATCCAGGAGGCGCGA 180
Db 121 GCGGCGCTCGCGGAACAAACCCGCTGCGGCAATGAGTCTCATCCAGGAGGCGCGA 180
QY 181 GAGGCTCGGCTATCAACGCGGATTTCAAGCGCGGCGCAACTGTCGCGACC 240
Db 181 GAGGCTCGGCTATCAACGCGGATTTCAAGCGCGGCGCAACTGTCGCGACC 240
QY 241 GCGGACGAGTGTGGGCGGAGCGCTGATTTATGCTCAAGGTCAAAAGAACGATAGCGGC 300
Db 241 GCGGACGAGTGTGGGCGGAGCGCTGATTTATGCTCAAGGTCAAAAGAACGATAGCGGC 300
QY 301 GAATACGCGCGCGTGGAGACGAGGAGATGTTGTCAGTTCGATTTGGCGCGCTCA 360
Db 301 GAATACGCGCGCGTGGAGACGAGGAGATGTTGTCAGTTCGATTTGGCGCGCTCA 360
QY 361 CGTCTTGAACCGATGCTGTTGATTCGCGACGACGTCATATGCTTACGAGACCGTC 420
Db 361 CGTCTTGAACCGATGCTGTTGATTCGCGACGACGTCATATGCTTACGAGACCGTC 420
QY 421 CAGACGCGAGCGGCGACTACCGCTGTCGCCGATGAGCGAGTGGCGGCGACTC 480
Db 421 CAGACGCGAGCGGCGACTACCGCTGTCGCCGATGAGCGAGTGGCGGCGACTC 480
QY 481 GCGGCGCGAGTGTGGCTTACCACTGATGCAACCAAGGGGCGCGGTGTCGATG 540
Db 481 GCGGCGCGAGTGTGGCTTACCACTGATGCAACCAAGGGGCGCGGTGTCGATG 540

```

```

QY 541 GCGGCGTGGCGCGCGTGAACCGGCGGAGCTGCTGATGATCGCGCGGCGACCGCGCG 600
Db 541 GCGGCGTGGCGCGCGTGAACCGGCGGAGCTGCTGATGATCGCGCGGCGACCGCGCG 600
QY 601 TACACGCGACCGCGCATTCGCGCAACGGCATGGGCGGACCGCTTACGATTCAC 660
Db 601 TACACGCGACCGCGCATTCGCGCAACGGCATGGGCGGACCGCTTACGATTCAC 660
QY 661 ATGACCAACTTGGCACTCGACCGCGAGTCTGCGGCGGATGACCAATGCTACTCA 720
Db 661 ATGACCAACTTGGCACTCGACCGCGAGTCTGCGGCGGATGACCAATGCTACTCA 720
QY 721 TCGGCTACGAGTTCGAGGGTGGCGTCAACGTCGCGACCTGATGATGGGCGCGCG 780
Db 721 TCGGCTACGAGTTCGAGGGTGGCGTCAACGTCGCGACCTGATGATGGGCGCGCG 780
QY 781 GTGCGAGCGCGCAAGGCGACCCCAATTAGTCTCGAATTCATTTGCGCGCATATGAACA 840
Db 781 GTGCGAGCGCGCAAGGCGACCCCAATTAGTCTCGAATTCATTTGCGCGCATATGAACA 840
QY 841 GTGCGGATGCTGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Db 841 GTGCGGATGCTGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 901 ACCACCTACGACACCGCGAGCTTGGCGGCGGATGATGATGATGATGATGATGATGAT 960
Db 901 ACCACCTACGACACCGCGAGCTTGGCGGCGGATGATGATGATGATGATGATGATGAT 960
QY 961 ATGCGGCTCTGCGGCGGAGAGCTGACCTGACGCTGACGCTGACGCTGACGCTGAT 1020
Db 961 ATGCGGCTCTGCGGCGGAGAGCTGACCTGACGCTGACGCTGACGCTGACGCTGAT 1020
QY 1021 GTGCTGAGCTTCCGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 1021 GTGCTGAGCTTCCGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1081 AAGGCTTTTCGACGACGAGGAGGCGTTACTGTCGGAAGGGGTGACCGACCTGGGG 1140
Db 1081 AAGGCTTTTCGACGACGAGGAGGCGTTACTGTCGGAAGGGGTGACCGACCTGGGG 1140
QY 1141 GTGCGCTTACCGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db 1141 GTGCGCTTACCGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
QY 1201 CACACNTCGGGAGTAANGAAGCGATGATGTCGNC 1235
Db 1201 CACACGTCGGGAGTAAGGAAGCGATGATGTCGCGC 1235

RESULT 6
AB9747 1237 bp DNA PAT 22-JAN-2000
LOCUS Sequence 4 from Patent WO9832862.
DEFINITION AB9747
ACCESSION AB9747
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1237)
AUTHORS Flohe L. and Singh M.
TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
JOURNAL Patent: WO 9832862-A 30-JUL-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
source 1..1237
Location/Qualifiers
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 236 a 394 c 386 g 221 t
ORIGIN

```

Query Match 99.6%; Score 1230.4; DB 5; Length 1237;
 Best Local Similarity 99.7%; Pred. No. 5,6e-169;
 Matches 1231; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 ATCTTGAGATTATGCAACTTTCTTACACTGAGGCTACAGTATGAGAGGGTATATC 60
    |||
DB 1 ATCTTGAGATTATGCAACTTTCTTACACTGAGGCTACAGTATGAGAGGGTATATC 60
QY 61 ATGGGGCTGGTATTCGACGAGACCAAAACAAAGATTCGGGGGCGCATCACCCCG 120
    |||
DB 61 ATGGGGCTGGTATTCGACGAGACCAAAACAAAGATTCGGGGGCGCATCACCCCG 120
QY 121 GCGGGCTCGCGAAGTAAACCGCTGCTGGCATAGAGTGTCTATCCAGGAGGTGCCGA 180
    |||
DB 121 GCGGGCTCGCGAAGTAAACCGCTGCTGGCATAGAGTGTCTATCCAGGAGGTGCCGA 180
QY 181 GAGGGCTCGGCTATACAGGAGGCGGATTCAGAGGCGGAGCGGCGCAACGTGCGGACC 240
    |||
DB 181 GAGGGCTCGGCTATACAGGAGGCGGATTCAGAGGCGGAGCGGCGCAACGTGCGGACC 240
QY 241 GCGGACGAGGTGCGGCGGAGCGGCTGATTTATGCTCAAGGTCAAAAGACCGATAGCGGC 300
    |||
DB 241 GCGGACGAGGTGCGGCGGAGCGGCTGATTTATGCTCAAGGTCAAAAGACCGATAGCGGC 300
QY 301 GAATAGCGGCGCTCGGACAGCGGAGATCTTGTTCAGGTTCTTGACATTTGCCCGCTCA 360
    |||
DB 301 GAATAGCGGCGCTCGGACAGCGGAGATCTTGTTCAGGTTCTTGACATTTGCCCGCTCA 360
QY 361 CGAGCTTGACAGCGATGCTGTTGGATTCGGGCGACGACGATTCGCTTCGCAACCGCTC 420
    |||
DB 361 CGAGCTTGACAGCGATGCTGTTGGATTCGGGCGACGACGATTCGCTTCGCAACCGCTC 420
QY 421 CAGACCGCGGCGGAGCGGAGCTACCTGCTGCTGCGGAGTGAAGCGAGCGGCTGCACTC 480
    |||
DB 421 CAGACCGCGGCGGAGCGGAGCTACCTGCTGCTGCGGAGTGAAGCGAGCGGCTGCACTC 480
QY 481 GCGGCGGAGTGGCGGCTTACACCTTATGCGAAGCCAAAGGGGCGCGGTGTGCTGATG 540
    |||
DB 481 GCGGCGGAGTGGCGGCTTACACCTTATGCGAAGCCAAAGGGGCGCGGTGTGCTGATG 540
QY 541 GGGGGGCTGGCGGCGCTGCAAGCGGCGGAGCGGCTGCTGATGCGGCGGCGGCGGCGG 600
    |||
DB 541 GGGGGGCTGGCGGCGCTGCAAGCGGCGGAGCGGCTGCTGATGCGGCGGCGGCGGCGG 600
QY 601 TACAGCGACCGCGCATCGCCAAAGGCGATGGCGCGAGCGCTTACGCTTACAGATCAAC 660
    |||
DB 601 TACAGCGACCGCGCATCGCCAAAGGCGATGGCGCGAGCGCTTACGCTTACAGATCAAC 660
QY 661 ATGACGACAAATTGGGCACTGCAAGCGGAGTTCGCGGGCGGATCCACACTCGCTACTCA 720
    |||
DB 661 ATGACGACAAATTGGGCACTGCAAGCGGAGTTCGCGGGCGGATCCACACTCGCTACTCA 720
QY 721 TCGGGCTACAGAGTGCAGGGTGCCTCAAAAGTSCCAAGCTGGGATGGGCGGCTGCTG 780
    |||
DB 721 TCGGGCTACAGAGTGCAGGGTGCCTCAAAAGTSCCAAGCTGGGATGGGCGGCTGCTG 780
QY 781 GTGCGAGGCGCCAAAGGCGCAATTAATGTCTCGAATTCACCTGTGCGGCAATGAAACCA 840
    |||
DB 781 GTGCGAGGCGCCAAAGGCGCAATTAATGTCTCGAATTCACCTGTGCGGCAATGAAACCA 840
QY 841 GGTGGCTACTGTGGATATAGCCATGAGACGAGGCGGCTGTTGCAAGGCTCACAGCCG 900
    |||
DB 841 GGTGGCTACTGTGGATATAGCCATGAGACGAGGCGGCTGTTGCAAGGCTCACAGCCG 900
QY 901 ACCAGCTACAGCACCGGAGCTTGCCTGACAGACAGCTGTTTACTCGTGGCGCAAC 960
    |||
DB 901 ACCAGCTACAGCACCGGAGCTTGCCTGACAGACAGCTGTTTACTCGTGGCGCAAC 960
QY 961 ATGCGGCGCTGCTGCGGAGGAGCTGACCTAGCGGCTGACCAAGCGGAGATGCCGTAT 1020
    |||
DB 961 ATGCGGCGCTGCTGCGGAGGAGCTGACCTAGCGGCTGACCAAGCGGAGATGCCGTAT 1020
QY 1021 GTGCTGAGAGCTTGCGGACGATGGCTGGCGGCGGCTGCGGCTGCAATCCGCGACTAGCC 1080
  
```

```

DB 1021 GTGCTGAGAGCTTGCGGACGATGGCTGGCGGCGGCTGCGGCTGCAATCCGCGACTAGCC 1080
    |||
QY 1081 AAGGCTCTTTCGACGACGAGGCGGCTTACTGTCTCCGAAGGAGTGGCCACGACCTGGGG 1140
    |||
DB 1081 AAGGCTCTTTCGACGACGAGGCGGCTTACTGTCTCCGAAGGAGTGGCCACGACCTGGGG 1140
QY 1141 GTGCGTACAGGAGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
    |||
DB 1141 GTGCGTACAGGAGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
QY 1201 CACACNTCGGAGTAANGAAGCATGATGTCGMC 1235
    |||
DB 1201 CACACNTCGGAGTAANGAAGCATGATGTCGMC 1235
    |||

RESULT 7
MTV002
LOCUS 56414 bp DNA BCT 17-JUN-1998
DEFINITION Mycobacterium tuberculosis H37Rv complete genome; segment 122/162.
ACCESSION AL008967 AL123456
VERSION AL008967.1 GI:3261491
KEYWORDS
SOURCE Mycobacterium tuberculosis.
ORGANISM Mycobacterium tuberculosis
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium.
REFERENCE 1 (bases 1 to 56414)
AUTHORS Cole,S.T., Brosch,R., Parkhill,J., Garnier,T., Churcher,C.,
Harris,D., Gordon,S.V., Eiglmeier,K., Gas,S., Barry III,C.E.,
Tekala,F., Badcock,K., Basham,D., Brown,D., Chillingworth,T.,
Connor,R., Davies,R., Devlin,K., Felwell,T., Gentles,S.,
Hamlin,N., Holroyd,S., Hornsby,T., Jagels,K., Krogh,A., McLean,J.,
Moule,S., Murphy,L., Oliver,S., Osborne,J., Quail,M.A.,
Rajandream,M.A., Rogers,J., Rutter,S., Seeger,K., Skelton,S.,
Squires,S., Squires,R., Sulston,J.E., Taylor,K., Whitehead,S. and
Barrell,B.G.
TITLE Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence
JOURNAL Nature 393 (6685), 537-544 (1998)
MEDLINE 98295987
REMARK 12:396(6707):1901]
REFERENCE 2 (bases 1 to 56414)
AUTHORS Parkhill,J.
TITLE Direct Submission
JOURNAL Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium
tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome
Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique
Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,
75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk
on Jun 27, 1998 this sequence version replaced gi:12642356.
COMMENT
NOTES:
Details of M. tuberculosis sequencing at the Sanger Centre are
available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/M_tuberculosis/) CDS have
been renumbered from the original cosmid submissions but the old
gene designations are in brackets after the new gene numbers.
Gene prediction was based on a Hidden Markov Model of TB genes
implemented in tRparse (Kiroh) supplemented with visual inspection
of positional base preference in codons, especially where there is
an increase in the observed/expected third position G + C.
CAUTION: In some cases we may not have predicted the correct
initiation codon. Where possible we choose an initiation codon
(atg, gtg, or ttg) which is preceded by an upstream ribosome
binding site sequence (optimally 5-11bp before the initiation
codon). If this cannot be identified we choose the most upstream
initiation codon.
FEATURES
source location/Qualifiers
1..56414
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
  
```


Query Match 99.6%; Score 1230.4; DB 1; Length 5644;
 Best Local Similarity 99.7%; Pred. No. 2.4e-169;
 Matches 1231; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1 ATCTTGAGATTATTCGAACTTCTTCACTGAAAGCGTACAGTACGAGAGGGTAAATC 60
 Db 38201 ATCTTGAGATTATTCGAACTTCTTCACTGAAAGCGTACAGTACGAGAGGGTAAATC 38260
 QY 61 ATGCGCGTTCGATTCGCAACCGAGACAAAACGAAATTCGGGTGGCCATCCGCCG 120
 Db 38261 ATGCGCGTTCGATTCGCAACCGAGACAAAACGAAATTCGGGTGGCCATCCGCCG 38320
 QY 121 GCGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 180
 Db 38321 GCGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 38380
 QY 181 GAGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 240
 Db 38381 GAGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 38440
 QY 241 GCGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 300
 Db 38441 GCGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 38500
 QY 301 GAATACGCGCGCTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 360
 Db 38501 GAATACGCGCGCTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 38560
 QY 361 GCGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 420
 Db 38561 GCGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 38620
 QY 421 CAGACCGCGCGCTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 480
 Db 38621 CAGACCGCGCGCTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 38680
 QY 481 GCGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 540
 Db 38681 GCGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 38740
 QY 541 GCGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 600
 Db 38741 GCGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 38800
 QY 601 TACACGCGCGCTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 660
 Db 38801 TACACGCGCGCTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 38860
 QY 661 ATGACGCGCGCTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 720
 Db 38861 ATGACGCGCGCTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 38920
 QY 721 TCGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 780
 Db 38921 TCGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 38980
 QY 781 GCGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 840
 Db 38981 GCGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 39040
 QY 841 GCGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 900
 Db 39041 GCGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 39100
 QY 901 AACACGCGCGCTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 960
 Db 39101 AACACGCGCGCTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 39160
 QY 961 ATGCGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 1020
 Db 39161 ATGCGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 39220
 QY 1021 GTGCTGAGCTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 1080

Db 39221 GTGCTGAGCTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 39280
 QY 1081 AAAGCTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 1140
 Db 39281 AAAGCTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 39340
 QY 1141 GTGCTGAGCTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 1200
 Db 39341 GTGCTGAGCTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 39400
 QY 1201 CACACNTGGGAGTAAAGAAAGGATGATGTCGNC 1235
 Db 39401 CACACNTGGGAGTAAAGAAAGGATGATGTCGNC 39435

RESULT 8

LOCUS A17608 1228 bp DNA PAT 22-JAN-2000
 DEFINITION Sequence 5 from Patent WO9836089.
 ACCESSION A87608
 VERSION A87608.1 GI:6736248
 KEYWORDS
 SOURCE
 ORGANISM
 unclassified
 unclassified
 unclassified

REFERENCE

1 (bases 1 to 1228)
 FLOHE, L. and Singh, M.
 TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
 JOURNAL Patent: WO 9836089-A 20-AUG-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source
 Location/Qualifiers
 1..1228
 /organism="unclassified"
 /db_xref="taxon:32644"

BASE COUNT 236 a 391 c 382 g 219 t
 ORIGIN

Query Match 99.1%; Score 1224.4; DB 5; Length 1228;
 Best Local Similarity 99.8%; Pred. No. 4.1e-168;
 Matches 1225; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 ATCTTGAGATTATTCGAACTTCTTCACTGAAAGCGTACAGTACGAGAGGGTAAATC 60
 Db 1 ATCTTGAGATTATTCGAACTTCTTCACTGAAAGCGTACAGTACGAGAGGGTAAATC 60
 QY 61 ATGCGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 120
 Db 61 ATGCGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 120
 QY 121 GCGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 180
 Db 121 GCGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 180
 QY 181 GAGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 240
 Db 181 GAGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 240
 QY 241 GCGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 300
 Db 241 GCGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 300
 QY 301 GAATACGCGCGCTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 360
 Db 301 GAATACGCGCGCTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 360
 QY 361 GCGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 420
 Db 361 GCGGCGTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 420
 QY 421 CAGACCGCGCGCTTCGCAACCGAGTACCGTGTGCTCATGAGTCTCATCAGGACAGTCCGGA 480


```

QY 961 ATGCCCCCTCGGAGGAGACGTGACCTACGCGCTGACCAACGAGCATGCGCTAT 1020
    |||||||
DB 961 ATGCCCCCTCGGAGGAGACGTGACCTACGCGCTGACCAACGAGCATGCGCTAT 1020
QY 1021 GTGCTGAGCTTGGCCAGCATGCGTGGGGGGGGGCGGCGGTGCAATCCGCACTAGCC 1080
    |||||||
DB 1021 GTGCTGAGCTTGGCCAGCATGCGTGGGGGGGGGCGGCGGTGCAATCCGCACTAGCC 1080
QY 1081 AAGAGTCTTTCGACGACGAGAGGGGCGTTACTCTCCGAAGGGGGGCGCACCGACCTGGGG 1140
    |||||||
DB 1081 AAGAGTCTTTCGACGACGAGAGGGGCGTTACTCTCCGAAGGGGGGCGCACCGACCTGGGG 1140
QY 1141 GTGCGCTTACCGAGCCCGCCAGCGTGTGCTGCTGACTCTGCGCGCTGTTACCGCGAG 1200
    |||||||
DB 1141 GTGCGCTTACCGAGCCCGCCAGCGTGTGCTGCTGACTCTGCGCGCTGTTACCGCGAG 1200
QY 1201 CACACNTGGGAGTANGAAGCATGA 1228
    |||||
DB 1201 CACACNTGGGAGTANGAAGCATGA 1228

RESULT 10
LOCUS A87609 1235 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 6 from Patent W09836089.
ACCESSION A87609
VERSION A87609.1 GI:6736249
KEYWORDS
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1235)
AUTHORS Flohe, L. and Singh, M.
TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL Patent: WO 9836089-A 20-AUG-1998;
FLOHE LEBOPOLD (DE); SINGH MANAVIR (DE)
FEATURES
    source 1..1235
    /organism="unclassified"
    /db_xref="taxon:32644"
BASE COUNT 235 a 395 c 384 g 220 t 1 others
ORIGIN
Query Match 98.7%; Score 1219; DB 5; Length 1235;
Best Local Similarity 99.6%; Pred. No. 2,5e-167;
Matches 1230; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

```

```

DB 360 CGTGTTCACACCATGCTGTGTGATTCGGGACACCATGCTCAATTCCTACGAGACCGTC 419
    |||||||
QY 421 CACACCCGCGAGCGGCGCTACTACCCCTGCTTGGCCCGCATAGAGAACTGCCCGGTGACTC 480
    |||||||
DB 420 CACACCCGCGAGCGGCGCTACTACCCCTGCTTGGCCCGCATAGAGAACTGCCCGGTGACTC 479
QY 481 GCGCCGAGGTTGGCGCTTACCACTGATGCGAACCAAGGGGGGCGGCGGTGCTGATG 540
    |||||||
DB 480 GCGCCGAGGTTGGCGCTTACCACTGATGCGAACCAAGGGGGGCGGCGGTGCTGATG 539
QY 541 GCGGGGGTCCCGCGCTCGAACCCGCGACGTCGTGTGATCGCGCGGCAACCGCGGC 600
    |||||||
DB 540 GCGGGGGTCCCGCGCTCGAACCCGCGACGTCGTGTGATCGCGCGGCAACCGCGGC 599
QY 601 TACAAGCGAGCCCGCATTCGCCAAGGCGCATGGGCGGCAACCGTTACAGGTTTACAGCAAC 660
    |||||||
DB 600 TACAAGCGAGCCCGCATTCGCCAAGGCGCATGGGCGGCAACCGTTACAGGTTTACAGCAAC 659
QY 661 ATGACAAACTTCGGCACTCGACGCGGAGTTCTGCGCGGATCCACACTGCTACTCA 720
    |||||||
DB 660 ATGACAAACTTCGGCACTCGACGCGGAGTTCTGCGCGGATCCACACTGCTACTCA 719
QY 721 TCGGCTTACGAGCTCGAGGCTGCGTCAACGTCGCGGACCTGGTATGCGCGCTCTG 780
    |||||||
DB 720 TCGGCTTACGAGCTCGAGGCTGCGTCAACGTCGCGGACCTGGTATGCGCGCTCTG 779
QY 781 GTCCAGGCGCCCAAGGCGACCCAAATAGTCTGCAATTCCTTCCGCAATAGAAACA 840
    |||||||
DB 780 GTCCAGGCGCCCAAGGCGACCCAAATAGTCTGCAATTCCTTCCGCAATAGAAACA 839
QY 841 GTGCGGTACTGTGTGATATAGCATGACGACGAGGGGCTGTTTGAAGGCTCAGACCG 900
    |||||||
DB 840 GTGCGGTACTGTGTGATATAGCATGACGAGGGGCTGTTTGAAGGCTCAGACCG 899
QY 901 ACCACCTACGACCCCGAGCTGCGCGGTGACGACACAGCTGTTTACTGCGGTGGGGAAC 960
    |||||||
DB 900 ACCACCTACGACCCCGAGCTGCGCGGTGACGACACAGCTGTTTACTGCGGTGGGGAAC 959
QY 961 ATCCCGCGCTGCGTGCAGAGAGCTGACCTGACGCTGACCAACGCGAGATGCGGTAT 1020
    |||||||
DB 960 ATCCCGCGCTGCGTGCAGAGAGCTGACCTGACGCTGACCAACGCGAGATGCGGTAT 1019
QY 1021 GTGCTGAGCTTCCGACCATGCTGCGGGGCGGCTGCGGTGCAATCCGACATGCGCATAGCC 1080
    |||||||
DB 1020 GTGCTGAGCTTCCGACCATGCTGCGGGGCGGCTGCGGTGCAATCCGACATAGCC 1079
QY 1081 AAGAGTCTTTCGACGACGACGAGGGGCTTACTCTCCGAAGGGGTGGCCACCGACCTGGGG 1140
    |||||||
DB 1080 AAGAGTCTTTCGACGACGACGAGGGGCTTACTCTCCGAAGGGGTGGCCACCGACCTGGGG 1139
QY 1141 GTGCGCTTACCGAGCCCGCCAGCGTGTGCTGACTCTGCGCGCTGTTACGCGCGAG 1200
    |||||||
DB 1140 GTGCGCTTACCGAGCCCGCCAGCGTGTGCTGACTCTGCGCGCTGTTACGCGCGAG 1199
QY 1201 CACACNTGGGAGTANGAAGCATGATGTCGNC 1235
    |||||
DB 1200 CACACNTGGGAGTANGAAGCATGATGTCGNC 1234

RESULT 11
LOCUS A89749 1235 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 6 from Patent W09832862.
ACCESSION A89749
VERSION A89749.1 GI:6738283
KEYWORDS
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1235)
AUTHORS Flohe, L. and Singh, M.
TITLE L-ALanine DEHYDROGENASE OF MYCOBACTERIUM MARITIMUM

```

JOURNAL Patent: WO 9832862-A 30-JUL-1998;

FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)

FEATURES Location/Qualifiers

source

1. 1235

/organism="unidentified"

/db_xref="taxon:32644"

BASE COUNT 235 a 395 c 384 g 220 t 1 others

ORIGIN

Query Match 98.7%; Score 1219; DB 5; Length 1235;

Best Local Similarity 99.6%; Pred. No. 2,5e-167;

Matches 1230; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1 ATCTGCAATTAATGAGACTTTCTTACACATGAGCGTACAGTATCGAGGGGTAATC 60
 DB 1 ATCTGCAATTAATGAGACTTTCTTACACATGAGCGTACAGTATCGAGGGGTAATC 60
 QY 61 ATGCGCGTGTATTCGACCGAGACCAAAACAGAAATTCGGGGTGC CATACACCCG 120
 DB 61 ATGCGCGTGTATTCGACCGAGACCAAAACAGAAATTCGGGGTGC CATACACCCG 120
 QY 121 GCGGCGTGTGCGGACTAAACCGCTGCTGGCCATGAGTGTCTATCCAGCAGTCCGGA 180
 DB 121 GCGGCGTGTGCGGACTAAACCGCTGCTGGCCATGAGTGTCTATCCAGCAGTCCGGA 180
 QY 181 GAGGGCTGCGTATACGACGCGGATTTCAAGGGGCGGCGGCACTGGTC GCACC 240
 DB 181 GAGGGCTGCGTATACGACGCGGATTTCAAGGGGCGGCGGCACTGGTC GCACC 240
 QY 241 GCGGACAGGTGTGGCCGACGCTGATTTATTCCTCAAGGTCAAAAGACGATAGCGGC 300
 DB 241 GCGGACAGGTGTGGCCGACGCTGATTTATTCCTCAAGGTCAAAAGACGATAGCGGC 300
 QY 301 GAATAGCGCGCTGTGGAACAGGGGAGATCTTTTACAGTCTTCTGATTTGGCGCGCTA 360
 DB 301 GAATAGCGCGCTGTGGAACAGGGGAGATCTTTTACAGTCTTCTGATTTGGCGCGCTA 360
 QY 361 CGATGTTGACCGATCGTGTGTGATTCGCGACACGATTCCTTACAGAGACGCTC 420
 DB 361 CGATGTTGACCGATCGTGTGTGATTCGCGACACGATTCCTTACAGAGACGCTC 420
 QY 421 CAGACCGCGGACGCGCTACCCCTGCTTGCCTCCGATGACGAAAGTCCGGCTGACTC 480
 DB 421 CAGACCGCGGACGCGCTACCCCTGCTTGCCTCCGATGACGAAAGTCCGGCTGACTC 480
 QY 480 GCGGCGGCGGCGCTTACCAACCTGATGCGCAACCAAGGGGCGCGGTGTGCTGATG 540
 DB 480 GCGGCGGCGGCGCTTACCAACCTGATGCGCAACCAAGGGGCGCGGTGTGCTGATG 540
 QY 541 GCGGCGGCGGCGCTTACCAACCTGATGCGCGGCGCGGCGGCGGCGGCGGCGGCGG 600
 DB 541 GCGGCGGCGGCGCTTACCAACCTGATGCGCGGCGCGGCGGCGGCGGCGGCGGCGG 600
 QY 601 TACACGACGCGCGATCGGCAACGCGCATGGGCGGACCGTTACGGTTCTAGACATCAAC 660
 DB 601 TACACGACGCGCGATCGGCAACGCGCATGGGCGGACCGTTACGGTTCTAGACATCAAC 660
 QY 661 ATGCAAGCGACCGCGATCGGCAACGCGCATGGGCGGACCGTTACGGTTCTAGACATCAAC 659
 DB 661 ATGCAAGCGACCGCGATCGGCAACGCGCATGGGCGGACCGTTACGGTTCTAGACATCAAC 659
 QY 721 TCGGCGTGTGCGGACTGCGGCTCAAAAGTGTGCGGACTGCGGCTGCGGCTGCGGCTG 780
 DB 721 TCGGCGTGTGCGGACTGCGGCTCAAAAGTGTGCGGACTGCGGCTGCGGCTGCGGCTG 780
 QY 781 GTGCGAGGCGCGACGCAACCAATTAATGATTCGATTCATTTGCGGCAATGAACA 840
 DB 781 GTGCGAGGCGCGACGCAACCAATTAATGATTCGATTCATTTGCGGCAATGAACA 840
 QY 841 GCGGCGTGTGCGGACTGCGGCTCAAAAGTGTGCGGACTGCGGCTGCGGCTGCGGCTG 900
 DB 841 GCGGCGTGTGCGGACTGCGGCTCAAAAGTGTGCGGACTGCGGCTGCGGCTGCGGCTG 900
 QY 901 ACCACTAGACACCGACCGTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
 DB 901 ACCACTAGACACCGACCGTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
 QY 961 ATGCGCGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
 DB 961 ATGCGCGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
 QY 1021 GTGCTGAGCTTCCGACCGATGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
 DB 1021 GTGCTGAGCTTCCGACCGATGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
 QY 1081 AAAGGCTTTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
 DB 1081 AAAGGCTTTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
 QY 1141 GTGCGGCTTACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
 DB 1141 GTGCGGCTTACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
 QY 1201 CACACNTGCGGAGTAAGGAGCGATGATGTCGNC 1235
 DB 1201 CACACNTGCGGAGTAAGGAGCGATGATGTCGNC 1235
 QY 1235 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1290
 DB 1235 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1290

RESULT 12

A87611

LOCUS

DEFINITION

SEQUENCE

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 1230; Conservative

0; Mismatches 4; Indels 1; Gaps 1;

98.7%; Score 1218.4; DB 5; Length 1235;

Pred. No. 3e-167;

/organism="unidentified"

/db_xref="taxon:32644"

BASE COUNT 236 a 394 c 385 g 220 t

ORIGIN

Query Match

Best Local Similarity

Matches 1230; Conservative

0; Mismatches 4; Indels 1; Gaps 1;

98.7%; Score 1218.4; DB 5; Length 1235;

Pred. No. 3e-167;

/organism="unidentified"

/db_xref="taxon:32644"

BASE COUNT 236 a 394 c 385 g 220 t

ORIGIN

Query Match

Best Local Similarity

Matches 1230; Conservative

0; Mismatches 4; Indels 1; Gaps 1;

98.7%; Score 1218.4; DB 5; Length 1235;

Pred. No. 3e-167;

/organism="unidentified"

/db_xref="taxon:32644"

BASE COUNT 236 a 394 c 385 g 220 t

ORIGIN

Query Match

Best Local Similarity

Matches 1230; Conservative

0; Mismatches 4; Indels 1; Gaps 1;

98.7%; Score 1218.4; DB 5; Length 1235;

Pred. No. 3e-167;

/organism="unidentified"

/db_xref="taxon:32644"

BASE COUNT 236 a 394 c 385 g 220 t

ORIGIN

Db 780 GTCCAGGCGCCAGGACCCAAATGATCTGTAATCTTGTCCGCAATATGAACA 839
QY 841 GGTGCGGACTGTGATATATGACATGACACGAGGCGGCTGTTTGAAGGCTCAGACCG 900
Db 840 GGTGCGGACTGTGATATATGACATGACACGAGGCGGCTGTTTGAAGGCTCAGACCG 899
QY 901 ACACACTACGACACCCGACGCTTGGCGTGACAGACACGCTGTTTACTGGCGGAC 960
Db 900 ACACACTACGACACCCGACGCTTGGCGTGACAGACACGCTGTTTACTGGCGGAC 959
QY 961 ATGCGCGCTGTGGTGGGGAAGACCTGACCTACCGCTGACACACGCGACATCCGAT 1020
Db 960 ATGCGCGCTGTGGTGGGGAAGACCTGACCTACCGCTGACACACGCGACATCCGAT 1019
QY 1021 GTGCTGAGCTTGGCGACATGCTGAGGCGGCGGCTGCGGTAATCCGCACTAGCC 1080
Db 1020 GTGCTGAGCTTGGCGACATGCTGAGGCGGCGGCTGCGGTAATCCGCACTAGCC 1079
QY 1081 AAGGCTTTTCGACGACGAGAGGGGCGTTACTGTCGAAAGGTCGACACCGACCTGGG 1140
Db 1080 AAGGCTTTTCGACGACGAGAGGGGCGTTACTGTCGAAAGGTCGACACCGACCTGGG 1139
QY 1141 GTGCGCTTACCGAGCCCGCCAGAGGTGCTGGCTGACTCTGCGCGCTCTGTTAGCGGAG 1200
Db 1140 GTGCGCTTACCGAGCCCGCCAGAGGTGCTGGCTGACTCTGCGCGCTCTGTTAGCGGAG 1199
QY 1201 CACACNTCGGAGTAANGAAGCATGATGTCGNC 1235
Db 1200 CACACNTCGGAGTAANGAAGCATGATGTCGNC 1234

RESULT 14
LOCUS A87605 1245 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 2 from Patent M09836089.
ACCESSION A87605
VERSION A87605.1 GI:6736245
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1245)
AUTHORS Flohe, L. and Singh, M.
TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL Patent: WO 9836089-A-20-AUG-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
source Location/Qualifiers
1..1245
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 238 a 398 c 387 g 222 t
ORIGIN

Query Match 98.5%; Score 1216; DB 5; Length 1245;
Best Local Similarity 99.3%; Pred. No. 6,7e-167;
Matches 1232; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

QY 1 ATCTTGAGATTATGCACTTTCTTCACTGAAAGCTACAGTATGAGAGGGGTATC 60
Db 1 ATCTTGAGATTATGCACTTTCTTCACTGAAAGCTACAGTATGAGAGGGGTATC 60
QY 61 ATGCGCGCTGTGGTGGGGAAGACCTGACCTACCGCTGACACACGCTGTTTACTGGCGGAC 114
Db 61 ATGCGCGCTGTGGTGGGGAAGACCTGACCTACCGCTGACACACGCTGTTTACTGGCGGAC 120
QY 115 ACCCGGCGCGGCTGCGGAACTAACCGTGCCTGAGAGTGTGCTCATCCAGCAGGT 174
Db 121 ACCCGGCGCGGCTGCGGAACTAACCGTGCCTGAGAGTGTGCTCATCCAGCAGGT 180
QY 175 GCGGGAAGGGCTGCTATACACGACGCGGATTTCAAGGCGGACGCGCAACTGGTC 234
Db 181 GCGGGAAGGGCTGCTATACACGACGCGGATTTCAAGGCGGACGCGCAACTGGTC 240

QY 235 GGCACCGCCGACAGGCTGTGGGCGGACGGCTGATTATGCTCAAGGTCGAAAGCCGATA 294
Db 241 GGCACCGCCGACAGGCTGTGGGCGGACGGCTGATTATGCTCAAGGTCGAAAGCCGATA 300
QY 295 GCGGCGGAATACGGCGGCTTGCACACGGGACAGATTTGTTACGTTTCTGATTTGGCC 354
Db 301 GCGGCGGAATACGGCGGCTTGCACACGGGACAGATTTGTTACGTTTCTGATTTGGCC 360
QY 355 GCGTACGTCCTTGCACCGATGCTGTTGGATTTCCGGACACGATCAATTGCTTACGAG 414
Db 361 GCGTACGTCCTTGCACCGATGCTGTTGGATTTCCGGACACGATCAATTGCTTACGAG 420
QY 415 ACCGTCAGACCGCCGACGCGGCTACCTACCGCTTGGCCCGGATGAGCGAAGTCGCGGT 474
Db 421 ACCGTCAGACCGCCGACGCGGCTACCTACCGCTTGGCCCGGATGAGCGAAGTCGCGGT 480
QY 475 CGACTGCGCCCGCAGGTTGGCGCTTACCACTGATGCGAACCAGGAGGCGCGGCTGTG 534
Db 481 CGACTGCGCCCGCAGGTTGGCGCTTACCACTGATGCGAACCAGGAGGCGCGGCTGTG 540
QY 535 CTGATGGGCGGGGTGCCCGGCTGCAACCGGCC GACGTGCTGTGATTCGGCGCGGACCC 594
Db 541 CTGATGGGCGGGGTGCCCGGCTGCAACCGGCCGACGTGCTGTGATTCGGCGCGGACCC 600
QY 595 GCGGCGTACAAAGCGGCGGCTATGCGCAAGGCGGCGGCGGCTTACGGTTCTAGAC 654
Db 601 GCGGCGTACAAAGCGGCGGCTATGCGCAAGGCGGCGGCGGCTTACGGTTCTAGAC 660
QY 655 ATCAACATCGACAACTTGGCACTCGACGCGGATTCGCGGCGGATTCACACTCGC 714
Db 661 ATCAACATCGACAACTTGGCACTCGACGCGGATTCGCGGCGGATTCACACTCGC 720
QY 715 TACTCATGCGCTTACGAGCTCGAGGTCGCTGCAACGTGCGACCTGTGATTTGGGGCC 774
Db 721 TACTCATGCGCTTACGAGCTCGAGGTCGCTGCAACGTGCGACCTGTGATTTGGGGCC 780
QY 775 GTCTGTCGCGGCGGCGCA GACCCAAATTAAGTCTCAATTCACTTCTCCGCAATAG 834
Db 781 GTCTGTCGCGGCGGCGCA GACCCAAATTAAGTCTCAATTCACTTCTCCGCAATAG 840
QY 835 AAACAGGTGCGGTACTGTGATATAGCATGACGAGGCGGCTGTTTCAAGGCTCA 894
Db 841 AAACAGGTGCGGTACTGTGATATAGCATGACGAGGCGGCTGTTTCAAGGCTCA 900
QY 895 CGACCGACCTTACGACACACCGGACGCTGCGGCTGACGACACGCTGTTTACTGCGTG 954
Db 901 CGACCGACCTTACGACACACCGGACGCTGCGGCTGACGACACGCTGTTTACTGCGTG 960
QY 955 GCGAATGCGCGGCTGCGGCGGAGAGCGTACCTAGCGCTGACCAACGCGAGATG 1014
Db 961 GCGAATGCGCGGCTGCGGCGGAGAGCGTACCTAGCGCTGACCAACGCGAGATG 1020
QY 1015 CCGTATGCTGCTGAGCTTGGCGACATGCTGGCGGCGGCGGCTGATTCGAGTCCGGA 1074
Db 1021 CCGTATGCTGCTGAGCTTGGCGACATGCTGGCGGCGGCGGCTGATTCGAGTCCGGA 1080
QY 1075 CTGCGCAAGGCTTTTCGACGACGAGGCGGCTTACTGTCCGAAGGCTGCGCACGAC 1134
Db 1081 CTGCGCAAGGCTTTTCGACGACGAGGCGGCTTACTGTCCGAAGGCTGCGCACGAC 1140
QY 1135 CTGGGGGTGCGCTTACCGGAGCCCGGACGAGTGTGCTGACTCTGCGGCGGCTGTTAC 1194
Db 1141 CTGGGGGTGCGCTTACCGGAGCCCGGCAAGGCTGCTGAGTGTGCTGAGTGTGCTGAG 1200
QY 1195 GCGGACACACNTCGGAGTAANGAAGCATGATGTCGNC 1235
Db 1201 GCGGACACACNTCGGAGTAANGAAGCATGATGTCGNC 1241

RESULT 15
LOCUS A89745 1245 bp DNA PAT 22-JAN-2000


```

DEFINITION      Sequence 2 from Patent WO9832862.
ACCESSION       A89745
VERSION         A89745.1 GI:6738279
KEYWORDS        .
SOURCE          unidentified.
ORGANISM        unidentified.
REFERENCE       1 (bases 1 to 1245)
AUTHORS        Flohe, L. and Singh, M.
TITLE          L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
JOURNAL        Patent: WO 9832862-A 30-JUL-1998;
               FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES       source
               1. 1245
               /organism="unidentified"
               /db_xref="taxon:3264"
BASE COUNT      238 a 398 c 387 g 222 t
ORIGIN
Query Match      98.5%; Score 1216; DB 5; Length 1245;
Best Local Similarity 99.3%; Pred. No. 6,7e-167;
Matches 1232; Conservative 0; Mismatches 3; Indels 6; Gaps 1;
QY      1 ATCTTGAGATTAAATCGAACTTTCTTCACTGAGCGTAC GATATGAGAGGGGTATC 60
DB      1 ATCTTGAGATTAAATCGAACTTTCTTCACTGAGCGTAC GATATGAGAGGGGTATC 60
QY      61 ATGGCGCGGTATTCGACGAGCAACCAAAACACG-----AATTCGGGTGGCCATC 114
DB      61 ATGGCGCGGTATTCGACGAGCAACCAAAACACG-----AATTCGGGTGGCCATC 120
QY      115 ACCCGGGCGCGGTGCGGGAATTAACCCGTGTCGTCATGAGTGTCTCATTCAGCGAGT 174
DB      121 ACCCGGGCGCGGTGCGGGAATTAACCCGTGTCGTCATGAGTGTCTCATTCAGCGAGT 180
QY      175 GCGGAGAGGGGTGCGGTATACGAGCGGATTTCAAGGGCGGCGGCAACTGTC 234
DB      181 GCGGAGAGGGGTGCGGTATACGAGCGGATTTCAAGGGCGGCGGCAACTGTC 240
QY      235 GGCACCGCGCAGCAGGTGTGGGCGGAGCGTGTATTATTGCTCAAGGTCAAAAGCAAGATA 294
DB      241 GGCACCGCGCAGCAGGTGTGGGCGGAGCGTGTATTATTGCTCAAGGTCAAAAGCAAGATA 300
QY      295 GCGGCGGAATACGGCGCGCTCGACAGCGGAGATCTTGTTCAGCTTCTTGCATTGGCC 354
DB      301 GCGGCGGAATACGGCGCGCTCGACAGCGGAGATCTTGTTCAGCTTCTTGCATTGGCC 360
QY      355 GCGTACGCTGCTGACGAGTGTGTTGGATTCCGACACAGTCATTAATGCTTACGAG 414
DB      361 GCGTACGCTGCTGACGAGTGTGTTGGATTCCGACACAGTCATTAATGCTTACGAG 420
QY      415 ACCGTCAGACCGCGGAGCGGCACTACCTGCTGCGCGGATGAGCAAGTCCCGGT 474
DB      421 ACCGTCAGACCGCGGAGCGGCACTACCTGCTGCGCGGATGAGCAAGTCCCGGT 480
QY      475 CGACTCGCGCGCGGAGCGGCTTACCACTGATGGAACCAAGGGGGCGGCGGTG 534
DB      481 CGACTCGCGCGCGGAGCGGCTTACCACTGATGGAACCAAGGGGGCGGCGGTG 540
QY      535 CTGATGGCGGGGTGCGGCGGTGCAACCGGCGGAGTGTGTGTATGGGCGGCGCACC 594
DB      541 CTGATGGCGGGGTGCGGCGGTGCAACCGGCGGAGTGTGTGTATGGGCGGCGCACC 600
QY      595 GCGGCGTCAACGAGCGCGCATGCGCAACGAGCATGGGCGGAGCGTTACGTTTACAG 654
DB      601 GCGGCGTCAACGAGCGCGCATGCGCAACGAGCATGGGCGGAGCGTTACGTTTACAG 660
QY      655 ATCAACATGCAAACTTGGCACTCGACGCGAGTTCTGGGCGCGGATCCACACTCGC 714
DB      661 ATCAACATGCAAACTTGGCACTCGACGCGAGTTCTGGGCGCGGATCCACACTCGC 720
QY      715 TACTCATGCGCTACGAGCTCGAGGGTGGCGTCAAAAGTCCGACCTGGTATGGGCGC 774

```

```

DB      721 TACTCATGCGCTACGAGCTCGAGGGTGGCGTCAAAAGTCCGAGCTGGTATGGGCGC 780
QY      775 GTCCGTGTGCGGCGGCGGCGCAAGCGCAATTAATGCTGATTTCACTGTGGCGCATATG 834
DB      781 GTCTGTGTGCGGCGGCGGCGCAAGCGCAATTAATGCTGATTTCACTGTGGCGCATATG 840
QY      835 AAACGAGTGTGCTGCTGATGATATGCAATGCAACGAGGCGGCTGTTTGAAGGCTCA 894
DB      841 AAACGAGTGTGCTGCTGATGATATGCAATGCAACGAGGCGGCTGTTTGAAGGCTCA 900
QY      895 CGACGAGCACTTACGACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 954
DB      901 CGACGAGCACTTACGACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
QY      955 GCGACATGCGCGCGCTGCGTGGTGGCGGAGAGGAGTGCAGCTACGCGCTGACCAAGCGAGATG 1014
DB      961 GCGACATGCGCGCGCTGCGTGGTGGCGGAGAGGAGTGCAGCTACGCGCTGACCAAGCGAGATG 1020
QY      1015 CCGTATGTGCTGAGCTTGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1074
DB      1021 CCGTATGTGCTGAGCTTGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
QY      1075 CTAGCCAAAGTCTTTCGACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1134
DB      1081 CTAGCCAAAGTCTTTCGACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
QY      1135 CTGGGGGTGCGGCTTACCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1194
DB      1141 CTGGGGGTGCGGCTTACCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
QY      1195 GCGGAGCACACNTGGGAGTAAAGGAGGATGATGTGNC 1235
DB      1201 GCGGAGCACACNTGGGAGTAAAGGAGGATGATGTGNC 1241

```

Search completed: June 22, 2000, 14:51:53
Job time: 17528 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 15:05:34 ; Search time 458.59 seconds
(without alignments)
673.777 Million cell updates/sec

Title: US-09-362-485-3

Sequence: 1 ATCTTGCGATTATTCGAC.....ANGAGCGATGATGTCGNC 1235

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311985 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq.36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1216	98.5	1245	1 V49626	Mycobacterium tube
2	1216	98.5	1260	1 V49610	Mycobacterium sp.
3	1216	98.5	1260	1 V49625	Mycobacterium tube
4	467.6	37.9	682	1 V49511	Mycobacterium mari
5	277.2	22.4	1125	1 N91423	Sequence of heat-r
6	212.4	17.2	2817	1 V52155	Streptococcus pneu
7	178.4	14.4	9280	1 V74442	Streptococcus aur
8	94.8	7.7	1074	1 T67971	H. pylori membrane
9	69.2	5.6	544	1 V43039	Streptococcus pneu
10	57.6	4.7	31096	1 V74370	Streptococcus pneu
11	50.8	4.1	390	1 Q21833	Randomising oligon
12	50.8	4.1	390	1 Q36859	PCR primer for 5'
13	50.4	4.1	1413	1 X34205	Mycobacterium spec
14	50.4	4.1	1722	1 X34206	Mycobacterium spec
15	50.4	4.1	13872	1 T68715	Streptomyces venez
16	49	4.0	1269	1 Q03065	Sequence encoding
17	47.6	3.9	3946	1 T93610	Mycobacterium tube
18	47.6	3.9	28598	1 T06769	Sorangium cellulos
19	47.6	3.9	28598	1 T89956	Mycobacterium spec
20	47.6	3.9	49377	1 V05287	The scrophen biosy
21	46.4	3.8	985	1 V44439	Mycobacterium tube
22	46.4	3.8	985	1 V45458	M. tuberculosis im
23	46	3.7	1620	1 Q22482	groEL-1 gene codin
24	46	3.7	2582	1 T73117	Actinoplanes sp. a
25	46	3.7	2668	1 Q22485	groEL-1 gene. Reco
26	46	3.7	17955	1 V56642	Actinoplanes sp. a
27	45.4	3.7	833	1 Q64203	snab gene encoding
28	45.4	3.7	5392	1 Q64201	Sequence compislin
29	42.4	3.4	1833	1 Q64206	snb gene encoding
30	42.4	3.4	2185	1 V84066	Clone P5-4 encodin
31	42.4	3.4	2186	1 Q99365	S. lividans protea
32	42.4	3.4	24379	1 T93095	Streptomyces fireo
33	42.4	3.4	24379	1 V25925	Streptomyces roseo
34	42.2	3.4	329	1 V44425	Mycobacterium tube

35	42.2	3.4	329	1 V64534	M. tuberculosis im
36	42	3.4	3946	1 T93610	Mycobacterium tube
37	42	3.4	4018	1 Q63879	Polyhydroxyalkanoa
38	41.8	3.4	882	1 V44403	Mycobacterium tube
39	41.8	3.4	882	1 V44512	M. tuberculosis im
40	41.6	3.4	12036	1 Q04668	FHA structural gen
41	41.2	3.3	2414	1 Q05926	Sequence encoding
42	41	3.3	29879	1 Q46806	eryA region of S.
43	40.8	3.3	2151	1 Q76252	HSV-2 protease, IC
44	40.8	3.3	2151	1 Q76261	HSV-2 protease/ICP
45	40.8	3.3	2472	1 Q84671	HSV-2 UL6 gene. N

ALIGNMENTS

RESULT 1	
V49626	1
ID V49626	standard; DNA; 1245 BP.
AC V49626;	
DT 20-NOV-1998	(first entry)
DE Mycobacterium tuberculosis	L-Alanine dehydrogenase 40 KD antigen.
KW SS: Alanine dehydrogenase;	tuberculosis; TB; Mycobacteria; L-alanine.
OS Mycobacterium tuberculosis.	
PN W09836089-R2.	
PD 20-AUG-1998.	
PF 29-JAN-1998; E00483.	
PR 29-JAN-1997; EP-1013338.	
PA (FLOH/) FLOHE L.	
PI Flohe L, Hutter B, Kolz A, Singh M;	
DR WPI: 98-457123/39.	tuberculosis from alanine dehydrogenase activity
PT - useful for, e.g. for diagnosis, differentiation of strains,	
PT monitoring vaccination and identification of mycobacterial	
PT inhibitors	
PS Disclosure: Fig 3.19; 55pp; German.	
CC The Mycobacterium tuberculosis L-alanine dehydrogenase (AlaDH) is used in	
CC the production of kits for diagnosing tuberculosis (TB) and other	
CC mycobacterial infections in humans or animals. Kits are used for direct	
CC diagnosis of TB on clinical samples (e.g. body fluids) and can	
CC differentiate between pathogenic and non-virulent strains, e.g. for	
CC identifying pathogens of the M. tuberculosis (M. t.) complex. The kit may	
CC also be used to identify substances that inhibit mycobacteria, for	
CC combating epidemics and for vaccination follow-up. Oligonucleotides	
CC derived from AlaDH are used similarly in diagnostic hybridisation tests,	
CC also for culture confirmation of isolated strains and for chromosome	
CC fingerprinting to detect/differentiate between mycobacteria, and for	
CC L-alanine-specific biotransformation reactions. AlaDH is an early	
CC antigen, present extracellularly after only a few days of growth, making	
CC it an ideal drug target.	
SQ Sequence 1245 BP; 238 A; 398 C; 387 G; 222 T;	
Query Match	98.5%; Score 1216; DB 1; Length 1245;
Best Local Similarity	99.3%; Pred. No. 1.2e-24;
Matches 1332; Conservative	0; Mismatches 3; Indels 6; Gaps 1;
QY 1	ATCTTGCAATTAATGAACTTTTCACATGAAAGCTACAGTTCAGAGGGTAATC 60
DB 1	ATCTTGCAATTAATGAACTTTTCACATGAAAGCTACAGTTCAGAGGGTAATC 60
QY 61	ATGCGCGTGGTATTCGACGACGACCAAAACACG-----AATTCGGGTGGCATC 114
DB 61	ATGCGCGTGGTATTCGACGACGACCAAAACACGAAATTCATTCGGGTGGCATC 120
QY 115	ACCCGCGCGCGCGTGGCGGAACTACCCGCTGCGGCATAGAGTCTTATCCAGGACGT 174
DB 121	ACCCGCGCGCGCGTGGCGGAACTACCCGCTGCGGCATAGAGTCTTATCCAGGACGT 180
QY 175	GCGGAGAGAGGCTGCGTATACGACGAGGATTCAGAGGCGGCGGCGCACTGGTC 234
DB 181	GCGGAGAGAGGCTGCGTATACGACGAGGATTCAGAGGCGGCGGCGCACTGGTC 240

```

QY 235 GGCACCGCCGACAGGTGTGGGCGGACGCTGATTTATCTCTCAAGSTCAAGAACCAGATA 294
DB 241 GGCACCGCCGACAGGTGTGGGCGGACGCTGATTTATCTCTCAAGSTCAAGAACCAGATA 300
QY 295 GCGGCGGAATAGGCGCGCTGCGACAGCGGAGATTTTTCACGTTCTTTCGATTTGGCC 354
DB 301 GCGGCGGAATAGGCGCGCTGCGACAGCGGAGATTTTTCACGTTCTTTCGATTTGGCC 360
QY 355 GGGTCAGTGTGACCGCATGGTGTGGATTCGGGAGACACGATTCGCTTCGAG 414
DB 361 GGGTCAGTGTGACCGCATGGTGTGGATTCGGGAGACACGATTCGCTTCGAG 420
QY 415 ACCGTCAGACCGCCGACGCGCAGTACCCTGCTTCTGCCCCGATGACGAGTCGCGGT 474
DB 421 ACCGTCAGACCGCCGACGCGCAGTACCCTGCTTCTGCCCCGATGACGAGTCGCGGT 480
QY 475 CGACTGCGCCCGCAGTGTGGCGTTACACCTGATTCGGAACCCAGGGGCGCGGTGTG 534
DB 481 CGACTGCGCCCGCAGTGTGGCGTTACACCTGATTCGGAACCCAGGGGCGCGGTGTG 540
QY 535 CTGATGGGCGGCGGCGCGGCGTGAACCGGCGGACGTCGTGGTGTGATCGGCGCGGAC 594
DB 541 CTGATGGGCGGCGGCGGCGGCGTGAACCGGCGGACGTCGTGGTGTGATCGGCGCGGAC 600
QY 595 GCGGCGTACAGCGACCGCGCATCGCCAAAGCGCATGGCGCGACGCTTACGATAGAC 654
DB 601 GCGGCGTACAGCGACCGCGCATCGCCAAAGCGCATGGCGCGACGCTTACGATAGAC 660
QY 655 ATCAACATGACACAAATTTGGGAACTGAGCGCGGATTCCTGGGCGCGGATTCACATGCG 714
DB 661 ATCAACATGACACAAATTTGGGAACTGAGCGCGGATTCCTGGGCGCGGATTCACATGCG 720
QY 715 TACTCATCGGCTTACAGAGTTCGAGGCTGCGGCTGACAAAGTGCAGCTGGTGTGAGGCGC 774
DB 721 TACTCATCGGCTTACAGAGTTCGAGGCTGCGGCTGACAAAGTGCAGCTGGTGTGAGGCGC 780
QY 775 GTCTGTGTCGACAGGCGCGCAAGGACCCAAATTAATGTCGAAATTCCTTGTGCGGCTATG 834
DB 781 GTCTGTGTCGACAGGCGCGCAAGGACCCAAATTAATGTCGAAATTCCTTGTGCGGCTATG 840
QY 835 AAACGAGTGGGCTACTGTGATATAGCATTCGACAGGCGGCGGCTTTCGAAGGCTCA 894
DB 841 AAACGAGTGGGCTACTGTGATATAGCATTCGACAGGCGGCGGCTTTCGAAGGCTCA 900
QY 895 CGACGACCACTACGACACCGGACGCTTGCCTGTGACAGACAGCTGTTTACTGCGGTG 954
DB 901 CGACGACCACTACGACACCGGACGCTTGCCTGTGACAGACAGCTGTTTACTGCGGTG 960
QY 955 GCGAATATCCCGGCTGCGTGCCTGCGAAGAGTGTGACCTACGCGCTGACCAACGCGAGATG 1014
DB 961 GCGAATATCCCGGCTGCGTGCCTGCGAAGAGTGTGACCTACGCGCTGACCAACGCGAGATG 1020
QY 1015 CCGTATGCTGCTGAGTGTGCGGACATGGCTGCGGCGGCGGCTGCGATTCGAGCA 1074
DB 1021 CCGTATGCTGCTGAGTGTGCGGACATGGCTGCGGCGGCGGCTGCGATTCGAGCA 1080
QY 1075 CTAGCCAAAGGCTTTTTCGACGACGACGAAGGCGGCTTACTGTTCGAAGCGGTGCGACCGAG 1134
DB 1081 CTAGCCAAAGGCTTTTTCGACGACGACGAAGGCGGCTTACTGTTCGAAGCGGTGCGACCGAG 1140
QY 1135 CTGGGGGTGCGCTTTCGACGACGCGCGGCGGCGGCTGCTGACTCTGCGCGCTGTTAC 1194
DB 1141 CTGGGGGTGCGCTTTCGACGACGCGCGGCGGCGGCTGCTGACTCTGCGCGCTGTTAC 1200
QY 1195 GCGGAGCACACNTCGGAGTAANGAAGGATGATGTCNC 1235
DB 1201 GCGGAGCACACNTCGGAGTAANGAAGGATGATGTCNC 1241

```

RESULT 2
 ID V49510 standard; DNA: 1260 BP.
 AC V49510;

DT 20-OCT-1998 (first entry)
 DE Mycobacterium sp. Aladh DNA.
 KM Alanine dehydrogenase; Aladh; ADH; diagnosis; tuberculosis; pathogen;
 KM swimmers disease; vaccine; epidemic; infection; identification; ss.
 OS Mycobacterium sp.
 PN M09832862-A2.
 PD 30-JUL-1998.
 PF 29-JAN-1998; E00484.
 PR 29-JAN-1997; EP-101339.
 PA (FLOH) FLOHE L.
 PI FLOHE L, Hutter B, Kolk A, Singh M;
 DR WPI: 98-427958/36.
 PT Nucleic acid encoding alanine dehydrogenase of Mycobacterium marinum
 PT - used for diagnosis of tuberculosis and other mycobacterial
 PT diseases, also for treatment and prevention, for drug screening and
 PT for bio-transformation
 PS Disclosure: Page 11: 57pp; German.
 CC This sequence encodes an alanine dehydrogenase (ADH) protein isolated
 CC from a Mycobacterium sp. This protein is used to diagnose tuberculosis
 CC and other mycobacterial infections (including 'swimmers' disease', caused
 CC by M. marinum, a fish pathogen) in humans or animals. The protein can
 CC also be used for control of epidemics and for vaccination, to screen for
 CC agents with anti-mycobacterial activity, and in bio-transformations that
 CC are specific for L-alanine. Also mycobacteria can be identified by
 CC analysis of genomic ADH sequences. ADH is an antigen that is secreted
 CC early during infection.
 SQ Sequence 1260 BP; 243 A; 403 C; 389 G; 225 T;

Query Match 98.5%; Score 1216; DB 1; Length 1260;
 Best Local Similarity 99.3%; Pred. No. 1,26-244;
 Matches 1232; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

```

QY 1 ATCTTGACGATTAATGACACTTCTTCACTGAAAGCTACAGTATCGAGGGGTATC 60
DB 16 ATCTTGACGATTAATGACACTTCTTCACTGAAAGCTACAGTATCGAGGGGTATC 75
QY 61 ATGCGCGTGGTATTCGACGAGACGACAAACAAACG-----AATTCGGGGTGGCCATC 114
DB 76 ATGCGCGTGGTATTCGACGAGACGACAAACAAACGATTCGAGGGGTGGCCATC 135
QY 115 ACCCGCGCGGCGTGGCGAACTAACCCGTGTGGCCATGAGGTGCTCATCGAGGACGT 174
DB 136 ACCCGCGCGGCGTGGCGAACTAACCCGTGTGGCCATGAGGTGCTCATCGAGGACGT 195
QY 175 GCGGAGAGGGGCTGCGCTATCAGCGACGCGGATTCGAAGCGGCGAGCGGCAACTGCTC 234
DB 196 GCGGAGAGGGGCTGCGCTATCAGCGACCGGATTCGAAGCGGCGAGCGGCAACTGCTC 255
QY 235 GGCACCGCCGACAGGTGTGGGCGGACGCTGATTTATGCTCAAGGTCAAGAACCAGATA 294
DB 256 GGCACCGCCGACAGGTGTGGGCGGACGCTGATTTATGCTCAAGGTCAAGAACCAGATA 315
QY 295 GCGGCGGAATACGGCGCGCTGCGACAGCGGCGAGATCTTGTACGTTCTTTCGATTTGGCC 354
DB 316 GCGGCGGAATACGGCGCGCTGCGACAGCGGCGAGATCTTGTACGTTCTTTCGATTTGGCC 375
QY 355 GGGTACGCTGTGACGAGTGGCTGTGGATTCGGGACGAGTCAATGGCTACGAG 414
DB 376 GGGTACGCTGTGACGAGTGGCTGTGGATTCGGGACGAGTCAATGGCTACGAG 435
QY 415 ACCGTCAGACCGCGGACGCGGCGACATCCCTGCTTGGCCGATGAGGAGAGTGGCGGT 474
DB 436 ACCGTCAGACCGCGGACGCGGCGGCACTACCTGCTTGGCCGATGAGGAGAGTGGCGGT 495
QY 475 CGACTGCGCCCGCAGTGTGGCGTTACCACTGATGCGAACCCAAAGGGGCGCGGTGTG 534
DB 496 CGACTGCGCCCGCAGTGTGGCGTTACCACTGATGCGAACCCAAAGGGGCGCGGTGTG 555
QY 535 CTGATGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 594
DB 556 CTGATGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 615

```

595 GCCGCTACAAACGACGCGCATCGCAAGGCGCGCGCGCGCTTACGGTTAGAC 654
 616 GCGGGCTACAAAGCAGCCCGCATCGCAAGGCGCGCGCGCGCTTACGGTTAGAC 675
 655 ATCAATCATGACAACTTGGGCACTCGACGCGGATTCGCGCGCGCGCATCGC 714
 676 ATCAATCATGACAACTTGGGCACTCGACGCGGATTCGCGCGCGCGCATCGC 735
 715 TACTCATCGGCTACGAGCTCGAGGCGCGCGCGCGCGCTTACGGTTAGAC 774
 736 TACTCATCGGCTACGAGCTCGAGGCGCGCGCGCGCGCTTACGGTTAGAC 795
 775 GTCTGTGTCGACGCGCGCAAGGCGCGCGCGCGCGCTTACGGTTAGAC 834
 796 GTCTGTGTCGACGCGCGCAAGGCGCGCGCGCGCGCTTACGGTTAGAC 855
 835 AAACGAGTGGTACTGCGGATATAGCATGACGCGGCGCGCGCGCTTACGGTTAGAC 894
 856 AAACGAGTGGTACTGCGGATATAGCATGACGCGGCGCGCGCGCTTACGGTTAGAC 915
 895 CGACGACGACCTACGACGACGCGCGCGCGCGCGCGCTTACGGTTAGAC 954
 916 CGACGACGACCTACGACGACGCGCGCGCGCGCGCGCTTACGGTTAGAC 975
 955 GCGAATATGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTACGGTTAGAC 1014
 976 GCGAATATGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTACGGTTAGAC 1035
 1015 CGGATATGCTGCGGCTTACGCGCGCGCGCGCGCGCGCGCTTACGGTTAGAC 1074
 1036 CGGATATGCTGCGGCTTACGCGCGCGCGCGCGCGCGCGCTTACGGTTAGAC 1095
 1075 CTAGCCAAAGGCTTTTGCAGCGACGAGAGGCGCGCGCGCTTACGGTTAGAC 1134
 1096 CTAGCCAAAGGCTTTTGCAGCGACGAGAGGCGCGCGCGCTTACGGTTAGAC 1155
 1135 CTGGGGGTCGCGCTTACGCGCGCGCGCGCGCGCGCGCGCTTACGGTTAGAC 1194
 1156 CTGGGGGTCGCGCTTACGCGCGCGCGCGCGCGCGCGCGCTTACGGTTAGAC 1215
 1195 GCCGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTACGGTTAGAC 1235
 1216 GCCGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTACGGTTAGAC 1255
 RESULT 3
 ID V49625
 AC V49625 standard; DNA; 1260 BP.
 DT 20-NOV-1998 (first entry)
 DE Mycobacterium tuberculosis
 KW ss: Alanine dehydrogenase; tuberculosis; TB; mycobacteria; L-alanine.
 OS Mycobacterium tuberculosis.
 PN M09836089-A2.
 PD 20-JAN-1998.
 PF 20-JAN-1998: E00483.
 PR 20-JAN-1997: EP-101338.
 PA (FLOH) FLOH L.
 PI Flohe L, Hutter B, Kolk A, Singh M;
 DR WPI; 98-457123/39.
 PT Diagnosis of, e.g. tuberculosis from alanine dehydrogenase activity
 PT useful for, e.g. for diagnosis, differentiation of strains,
 PT monitoring vaccination and identification of mycobacterial
 PT inhibitors
 PS Claim 13: Fig 2.3: 55pp: German.
 CC The Mycobacterium tuberculosis alanine dehydrogenase (AlaDH) is in
 CC the production of kits for diagnosing tuberculosis (TB) and other
 CC mycobacterial infections in humans or animals. Kits are used for direct
 CC diagnosis of TB on clinical samples (e.g. body fluids) and can
 CC differentiate between pathogenic and non-virulent strains, e.g. for
 CC identifying pathogens of the M. tuberculosis (M. t.) complex. The kit may
 CC also be used to identify substances that inhibit mycobacteria, for
 CC combatting epidemics and for vaccination follow-up. Oligonucleotides

CC derived from AlaDH are used similarly in diagnostic hybridisation tests,
 CC also for culture confirmation of isolated strains and for chromosome
 CC fingerprinting to detect/differentiate between mycobacteria, and for
 CC L-alanine-specific biotransformation reactions. AlaDH is an early
 CC antigen, present extracellularly after only a few days of growth, making
 CC it an ideal drug target.
 SQ Sequence 1260 BP; 243 A; 403 C; 389 G; 225 T;
 Query Match 98.5%; Score 1216; DB 1; Length 1260;
 Best Local Similarity 99.3%; Pred. No. 1.2e-244;
 Matches 1232; Conservative 0; Mismatches 3; Indels 6; Gaps 1;
 1 ATCTTGACATTAATGAACTTCTTCACTGAGCGCTACAGATGAGAGGGGTAATC 60
 16 ATCTTGACATTAATGAACTTCTTCACTGAGCGCTACAGATGAGAGGGGTAATC 75
 61 ATGCGGTGGTATTCCGACGAGCAAAACAAACG-----AATCGGGTGGCCATC 114
 76 ATGCGGTGGTATTCCGACGAGCAAAACAAACGAAATTCGAATTCGGGTGGCCATC 135
 115 ACCCGGCGCGCGCGCGGAACTAACCGCTGCGCATGAGTGTCTATCCAGGAGGT 174
 136 ACCCGGCGCGCGCGCGGAACTAACCGCTGCGCATGAGTGTCTATCCAGGAGGT 195
 175 GCGGAGAGGCGCTCGGCTATCAACGAGCGGATTTCAAGGCGCGCGCGCTGCTC 234
 196 GCGGAGAGGCGCTCGGCTATCAACGAGCGGATTTCAAGGCGCGCGCGCTGCTC 255
 235 GGCACGCGCGCGCGCGGAGTGGGCGCGCGCGCGCTGATTTATGCTCAAGGTCAAGAGCGATA 294
 256 GGCACGCGCGCGCGCGGAGTGGGCGCGCGCGCGCTGATTTATGCTCAAGGTCAAGAGCGATA 315
 295 GCGGCGGAAATAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTACGGTTAGAC 354
 316 GCGGCGGAAATAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTACGGTTAGAC 375
 355 GCGTACGCTGCTTGCACCG. GCGTGTGATTCGCGCGCGCGCGCGCGCTTACGGTTAGAC 414
 376 GCGTACGCTGCTTGCACCGATCGTGTGATTCGCGCGCGCGCGCGCTTACGGTTAGAC 435
 415 ACCGTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTACGGTTAGAC 474
 436 ACCGTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTACGGTTAGAC 495
 475 CGACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTACGGTTAGAC 534
 496 CGACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTACGGTTAGAC 555
 535 CTGATGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTACGGTTAGAC 594
 556 CTGATGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTACGGTTAGAC 615
 595 GCCGCTACAAACGACGCGCGCATCGCAAGGCGCGCGCGCGCGCTTACGGTTAGAC 654
 616 GCCGCTACAAACGACGCGCGCATCGCAAGGCGCGCGCGCGCGCTTACGGTTAGAC 675
 655 ATCAATCATGACAACTTGGGCACTCGACGCGGATTCGCGCGCGCGCATCGC 714
 676 ATCAATCATGACAACTTGGGCACTCGACGCGGATTCGCGCGCGCGCATCGC 735
 715 TACTCATCGGCTACGAGCTCGAGGCGCGCGCGCGCGCGCTTACGGTTAGAC 774
 736 TACTCATCGGCTACGAGCTCGAGGCGCGCGCGCGCGCGCTTACGGTTAGAC 795
 775 GTCTGTGTCGACGCGCGCAAGGCGCGCGCGCGCGCGCTTACGGTTAGAC 834
 796 GTCTGTGTCGACGCGCGCAAGGCGCGCGCGCGCGCGCTTACGGTTAGAC 855
 835 AAACGAGTGGTACTGCGGATATAGCATGACGCGGCGCGCGCGCTTACGGTTAGAC 894
 856 AAACGAGTGGTACTGCGGATATAGCATGACGCGGCGCGCGCGCTTACGGTTAGAC 915

Qy	895	GCACCAGCACCTCAGACACACCCGAGTTGGCGTACAGAACCTGTTTACAGCCTG	954
Db	916	GCACGACACACTACAGACACACCGAGTTGGCGTACAGAACCTGTTTACAGCCTG	975
Qy	935	GCGAACATGCCCCCGCTGGTGGCCGGAAGACGTGACCTACAGCGCTGACCAAGCAGCATG	1014
Db	976	GCGAACATGCCCCCGCTGGTGGCCGGAAGACCTCAGCTTACGCGCTATACCAACCGACGATG	1035
Qy	1015	CCGATATGTGCTCGAGCTTGTGGCGACATGAGTGGCGGCGCGCTGGCGGTGCAATCCGACA	1074
Db	1036	CCGATATGTGCTCGAGCTTGTGGCGACATGAGTGGCGGCGCGCTGGCGGTGCAATCCGACA	1095
Qy	1075	CTAGCCAAAGGTCCTTGTGAGCGACGACGAAGGGCGCTTACTCCGAAGGCTGGCCACCGAC	1134
Db	1096	CTAGCCAAAGGTCCTTGTGAGCGACGACGAAGGGCGCTTACTCCGAAGGCTGGCCACCGAC	1155
Qy	1135	CTGGGGGTGCGCTTCACCGAGCCCGCAGCGTCTGAGCTTCGAGCCGCTGTTTAC	1194
Db	1156	CTGGGGGTGCGCTTCACCGAGCCCGCAGCGTCTGAGCTTCGAGCCGCTGTTTAC	1215
Qy	1195	GCCGAGCACACNTGCGGAGTAGAANGAAAGCATATGTCNC	1235
Db	1216	GCCGAGCACACGTCGGAGTAGAANGAAAGCATATGTCGC	1256

RESULT	4
V49511	
ID	V49511 standard; DNA; 682 BP.
NC	W49511

DT	20-OCT-1998 (first entry)
DE	Mycobacterium marinum Mar3 DNA.
KW	Alanine dehydrogenase; ALADH; ADH; diagnosis: tuberculosis; pathogen
KW	Swimmers disease; vaccine; epidemic; infection; identification; ss.
OS	Mycobacterium marinum.
PH	key
FT	location/qualifiers
FT	1 682
FT	/*tag_a
FT	/codon_start-3
FT	/product- "Mar3"
FT	/note- "Alanine dehydrogenase"

PN MO3832862--A2.
PD 30-JUL-1998.
PF 29-JAN-1998: E00484.
PR 29-JAN-1997: EP-101339.
PA (FLOH/) FLOHE L.
P1 Flohe L, Hutter B, Kolz A, Singh M;
DR WPI; 98-427958/36.
P-PSDB: W6t481.
PT Nucleic acid encoding alanine dehydrogenase of *Mycobacterium marinum*
used for diagnosis of tuberculosis and other mycobacterial
diseases, also for treatment and prevention, for drug screening and
for bio-transformation
P5 Claim 1; Page 34-35; 37pp; German.
CC This sequence encodes an alanine dehydrogenase (ADH) protein, Mar3
isolated from *Mycobacterium marinum*. This protein is used to diagnose
tuberculosis and other mycobacterial infections (including 'swimmers'
disease', caused by *M. marinum*, a fish pathogen) in humans or animals.
CC The protein can also be used for control of epidemics and for
vaccination, to screen for agents with anti-mycobacterial activity, and
for L-transformation, to screen for agents with anti-mycobacterial activity, and
can be identified by analysis of genomic ADH sequences. ADH is an antigen
that is secreted early during infection.
SQ Sequence 682 Bp; 105 A; 254 C; 225 G; 98 T;

Query Match	37.9%	Score 467.6	DB 1	length 682;
Best Local Similarity	80.4%	Pred. No. 3.9e-89		
Matches 548; conservative	0;	Mismatches 134;	Indels 0;	Gaps 0
Qy	98	AATTCCGGGTGCACATCACCCGGCGCGGTGCGGAACT	^CCCCGTGTGCCATGAG	157
Db	1	AATTCGGGTACGATCACCCGGCGCGGTGCGCCCTTGACCAACGCGCGCACAG	60	

OY	158	TGTCATCATCAGGAGAGTCCGAGAAAGGCGCTGGCTATACACGACCGGATTTCAAAAGCGG	217
Db	61	TGCTGATCATCAGGCGGATCCCGGAAAGGCTCTCGCATCTCCGACGCGGACTTCAAGGCTCG	120
OY	218	CAGCGCGCAACTGATGTCGCGACCGCCGCAAGGTGTGGGCGGACGCTGATTTATTGCTCA	277
Db	121	CCGGTGGCCACTGATGCATGACGCCCGCCGCAAGGTGTGGGCGGAGTGCAGACTGCTGCTCA	180
OY	278	AGGTCAAAAGAACCCGATTACCGCGGGAATATACGGCGCGCTGCGACACGGGCGAGATCTGTGTCA	337
Db	181	AGGTCAAGAAACCGATCGATGCTCAGATGACGCGCGGCTGTGCGCGGCGCCAGACCTGTTC	240
OY	338	CGTTCTTTCATTTGGCCGCGTTCACGCTGTGCACGATGCGTGTGGATTCGGGCACCA	397
Db	241	CCTACTCTCAGCTTGGCCCGC CGGCGCCCTGACACGATGCTCTCTTAAGTCCGGCACCA	300
OY	398	CGTCATATTCCTTAGAGACCGCTGCAGACCGCGCGACGCGCGACTACCCCTGCTTCCCGGA	457
Db	301	CGTCATCGCTTACGAGACGGGTGAGACCGCGCGAGCGGCGATTCGCTGTGTGGCCCCCA	360
OY	458	TGAGGAGATGGCGCGGTGAGATCCGCCGCCCAAGGTTGGCGCTTACCACCTGATGGAAACC	517
Db	361	TGAGGAGATGGCGCGGGGCGCTGTCTCCGCCCAAGCTGGGGGCTTACCACCTGATGGCACCC	420
OY	518	AAGGGGCGCGGATGTGCTGATAGCGCGGGGTGCCCGGCGTGAACCGCGGACGCTCTGTG	577
Db	421	ACGGCGGTTCGCGCGGTGCTGATGAGCGCGGTGCCCGGCGTCAACCTTCGCGACGTGTGG	480
OY	578	TGATCGGGGCGCGGACCGCGCGGTACAAACGAGGCGCGCATGCCCAAGGGCATGGGCGCGGA	637
Db	481	TGATCGGGGCGCGGACCGCGCGGATACAAACGCGCGCGCGCTGTGCCCAAGGGCATGGGCGCGGA	540
OY	638	CCGTTACGCTTCTAGACATCAATCGACAACAACTTGGGCAACTGCGAGCTGTGCG	697
Db	541	TGGTACCGCTCTGGATGTCAACATCAACAAGCTCCCGCATGACGAGCGGAGTTGCGCG	600
OY	698	GCCGGATACACATCTCGTACTCATCGGCGCTTCAGAGCTGGAAGGTGCGCGTAAAGTCCCG	757
Db	601	GTCGGGTCCGGAACCCGGTACTGCTGCACCCTCGACCTCTGAAGATGCGGCGAGTCCACCCCG	660
OY	758	ACCTGTGATTTGGGCGCGTCT 779	
Db	661	ACATGTGATTTGGGCGCGTCT 682	

RESULT	5
N91423	
ID	N91423 standard; DNA; 1125 bp

DT	01-FEB-1991	(first entry)
DE	Sequence of heat-resistant alanine dehydrogenase (AH) gene with mol. wt.	
DE	2 MD or less	
KW	Enzyme: ds.	
OS	Bacillus stearothermophilus IFO 12550.	
PN	J01043194-A.	
PD	15-FEB-1989.	
PR	10-AUG-1987; 200524.	
PR	10-AUG-1987; JP-200524.	
PA	(NIRA) Unilika KK.	
DR	WPI: 89-096096/13.	
PT	Recombinant plasmid for transforming Escherichia coli -	
PT	obtd. by connecting heat resistant alanine dehydrogenase gene to	
PT	vector plasmid	
PS	Disclosure: Fig 2, p693 : 9pp; Japanese.	
CC	A recombinant plasmid contg heat-resistant alanine dehydrogenase (AH)	
CC	gene with mol. wt. 2 or less. MD is claimed. Cells transformed with the	
CC	vector produce high levels of heat-resistant AH. transformed E. coli	
CC	is useful as a clinical inspection reagent.	
SC	Sequence 1125 BP; 246 A; 295 C; 367 G; 217 T;	

Query Match	22.48;	Score 277.2;	DB 1;	Length 1125;
Best Local Similarity	56.28;	Pred. No. 1.5e-49;		

Matches 604; Conservative 0; Mismatches 458; Indels 12; Gaps 4;

```

OY 61 ATGCCGCTGCTATTCACGACGACCAAAACAAAGATTCGGGTGCGCATCCCG 120
Db 1 ATGAAGATCGGCTATTCACAAAGAAATCAAAATCAAAATCGGCTGCGCATCC 60
OY 121 GCGGCGCTGCGGACCTAACCCGCTGAGAGTGTCT---CATCAGGACAGTGC 177
Db 61 GCGGCGCTGAGTACCTGCTCAAAACGGGCGCTAGAGTGTGTATGTGACGAAAGC 120
OY 178 GAGAGG-CTGCGCTATACCGACGCGGATTTCAAGCGCGGACGCGCAACTGTGCGC 237
Db 121 GCGGCTGAGTGTGCTTTCGATTCGAGTGTGAAAAAGCGGCGGAGTGTGCTGC 180
OY 238 ACCGCGACAGAGTGTGGGCGGACGCTATTTATGCTCAAGTCAAAAGACGATACG 297
Db 181 CGAAGCTGAGAGATGCTTGGACGCGGAGATGTTGAAAGTGAAGAGCCGCTGCT 240
OY 298 GCGAATACGCGCGCTGCGACACGCGGACATCTTGTCACTTCTGCTATTTGCGCGC 357
Db 241 CGAGAGTTCGCTATTTTCGCGCGGATTTGATTTGTATTTGATTTGATTTGATTTG 300
OY 358 TCACGTGCTTGCACGATGCTGTTGGATTCGCGACACGCTCAATTGCTACGAGAC 417
Db 301 GCGGACGCTGACGAAAGCCGCTGCTGAGCAAAAGTGTGCGGCTGCTACGAGAC 360
OY 418 GTCGACGCGCGGACGCGGACCTACCTGCTGCGCGGATGAGGACGCGGCTGGA 477
Db 361 GTCGACGCTGCGGACGCGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
OY 478 CTCGCGCGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 537
Db 421 ATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
OY 538 ATGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 597
Db 481 CTCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
OY 598 GCGTACACGCGGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 657
Db 541 GCGGACGCGGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
OY 658 AACATCGAACAATCTGCGCAACTGACGCGGATTTCTGCGCGGATTTCTGCTG 717
Db 601 AACGCGGCGGCTGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
OY 718 TCATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 777
Db 661 TCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
OY 778 CTGCTGCGGCGGCGGACGCGGACGCGGATTTCTGCTGCTGCTGCTGCTGCTG 837
Db 721 TTGATCCGCGGCGGCGGAA---AGCGAAGCTGTGAGGAGAAAGATGTGCTGCTG 777
OY 838 CCAGGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 894
Db 778 CCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 837
OY 895 GCGGCGGCGGCGGACGCGGACGCGGATTTCTGCTGCTGCTGCTGCTGCTGCTG 954
Db 838 CGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 897
OY 955 GCGAATACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1014
Db 898 GCGAATACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 954
OY 1015 CCGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1074
Db 955 CCGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1014
OY 1075 CTGACGCAAGGCTCTTTCGACGACGAGGCGGCTTACTGCTGCAAGCGGCTG 1128
Db 1015 CTGTTAAAGGATTCACACGCTGCGGCGGACATCTGCTGACGAAAGCGGCTG 1068

```

```

RESULT 6
ID V52155/c
V52155 standard; DNA; 28171 BP.
AC V52155;
AT 23-Oct-1998 (first entry)
DE Streptococcus pneumoniae genome fragment SEQ ID NO:22.
KW Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
OS Streptococcus pneumoniae; vaccine; pharmaceutical composition; ds.
PN MO9818931-A2.
PD 07-MAY-1998.
PF 30-Oct-1997; U19588.
PR 31-Oct-1996; US-028960.
PI (HUMA-) HUMAN GENOME SCI INC.
PI Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M,
PI Kunsch CA, Rosen CA.
DR WPI: 98-272225/24.
PT Computer-readable medium with recorded Streptococcus pneumoniae
PT polynucleotide sequences - useful in diagnostic kits and assays, and
PT pharmaceutical compositions and vaccines for Streptococcus
PT pneumoniae
PS Claim 1: Page 273-289; 1409bp; English.
CC The present invention describes a computer readable medium which has
CC the nucleotide sequences SEQ ID NO:1 to 391 (V52134 to V52524) recorded
CC on it or a representative fragment or a sequence at least 95% identical
CC to SEQ ID NO:1 to 391. The nucleotide sequences depicted in SEQ ID NO:1
CC to 391 (V52134 to V52524) are genomic fragments from Streptococcus
CC pneumoniae. The present invention also describes an isolated nucleic acid
CC molecule encoding a homologue of any of the fragments of the S.pneumoniae
CC genome (SEQ ID NO:1 to 391) where the nucleic acid molecule is produced
CC by a process comprising: (a) screening a genomic DNA library using as a
CC probe a target sequence defined by any of the sequences in SEQ ID NO:1
CC to 391, identifying members of the library which contain sequences
CC that hybridize to the target sequence and isolating the nucleic acid
CC molecules from the members; or (b) isolating mRNA, DNA or cDNA produced
CC from an organism, amplifying nucleic acid molecules whose nucleotide
CC sequence is homologous to amplification primers derived from the
CC fragment of the S.pneumoniae genome to prime the amplification and
CC isolating the amplified sequences. The computer readable medium can be
CC used in a computer-based system for identifying fragments of the
CC S. pneumoniae genome of commercial importance, or expression modulating
CC fragments of the S. pneumoniae genome. Products from the present
CC invention can be used in diagnosis kits and assays, and pharmaceutical
CC compositions and vaccines for S. pneumoniae.
SQ Sequence 28171 BP; 8507 A; 5080 C; 6440 G; 8144 T;

Query Match 17.2%; Score 212.4; DB 1; Length 28171;
Best Local Similarity 53.9%; Pred. No. 5,7e-36;
Matches 533; Conservative 0; Mismatches 436; Indels 19; Gaps 4;

OY 61 ATGCCGCTGCTATTCACGACGACCAAAACAAAGATTCGGGTGCGCATCCCG 120
Db 21711 ATGTTATCGGAATCCCAAAAGAAATTAATAAGCAAAACGCTGCGCTCACACT 21652
OY 121 GCGGCGCTGCGGGAACCTAACCCGCTGAGAGTGTCTGCTGCTGCTGCTGCTG 180
Db 21651 GCGAGTGTCTATGACTTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 21592
OY 181 GAGGCTGCGCTATCACGACGCGGATTTCAAGCGCGGACGCGCAACTGTGCGCAC 240
Db 21591 CTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 21532
OY 241 GCGGACGAGGCTGCGGCGGCGGATTTATGCTCAAGTCAAAAGACGATAGGCGG 300
Db 21531 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 21475
OY 301 GAATACGCGCGCTGCGACACGCGGAGATTTGCTGCTGCTGCTGCTGCTGCTG 360
Db 21474 GAATACGCTTACTTGGCGGACGATCTTCTTCTTACCTTGTGCAATAGCGCGCTGCT 21415

```

```

QY 361 GGTCTTCACCGATGCTGTGATTCGCGACCGCTCAATTCCTACGAGACCGTC 420
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 21414 CAGGATTAAGCAGATGATGTT-----AACAGCAAAAACAACTGAATCTT 21367
QY 421 CAGACCGCGAGGCGGCGATACCGCTGCTGCGCGATGAGAGAGTCCGCGTGCATC 480
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 21366 CGTGCAATCAAGGACAACTACCGCTGCTGCTTCTATGAGAGTGCAGGTGTATG 21307
QY 481 GCCGCCAGAGTGTGGCTTACCACTGATGCGAACCCAGGGGCGCGGTGTGTATG 540
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 21306 GCTGTTAATGAGAGCTACCTTCTTACTAGCAAGGTGGGTCTGTGTGTCTACTT 21247
QY 541 GCGGGGGTCCCGGCGTGAACCGCGACGTCGTGTGTGTGTGTGTGTGTGTGTGT 600
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 21246 GGTGTGTACACAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 21187
QY 601 TACAACGAGCGCGCGCGATGCGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 21186 ACAGATGCTGCGCGCGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 21127
QY 661 ATGCAAACTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 21126 TCCAGCGCTGCTCAAGTCTAGAACAGTCTTGGAGTCAAACTCTTATGTCT 21067
QY 721 TCGGCGCTACGAGCTGAGGCGTCCGTCACCAAGCGCGCGCGCGCGCGCGCGCG 780
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 21066 AATTTCATCAACATTTGAAGCAAGTGTGAGATGTGTGTGTGTGTGTGTGTGT 21007
QY 781 GTGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 21006 ATCCCTGCTGCAAAAGCACCGGATGCTGACAGATGATGTGTGTGTGTGTGT 20947
QY 841 GGTGCGCTACGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 897
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 20946 GCGTCTGTA-TCGTGAGCTGTCTTACCAAGTGTGTGTGTGTGTGTGTGTGT 20888
QY 898 CGGACCACTACGAGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 957
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 20887 GTGACAAAGCAGATGAGACCGGTCTATGAAAAACAGGTGTGTGTGTGTGTGT 20828
QY 958 AACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1017
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 20827 AATATCCCTGCTGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 20768
QY 1018 TATGTGCTGAGCTTCCGCGACCATGGCT 1045
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 20767 TATATCGAAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 20740

```

```

FT misc-feature 4441..4500
FT /tag-c
FT /note- "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT misc-feature 6241..6300
FT /tag-d
FT /note- "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT misc-feature 8041..8100
FT /tag-e
FT /note- "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"

```

Query Match 14.4%; Score 178.4; DB 1; Length 9280;

Best Local Similarity 50.6%; Pred. No. 6.1e-29; Matches 484; Conservative 0; Mismatches 466; Indels 6; Gaps 2;

```

QY 151 CAGGAGTGTCTTCACCGACGCTGCGGAGAGTGTGGGATATCCAGACCGGATTC 210
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 9280 CATACGTTTATGAGAACAAATGCGGTTTCAGATCTTTGAAAGATGATGATG 9221
QY 211 AAGCGGCGGCGCGCACTGTGCGACCGCGCGAGAGTGTGGCGACGCTATTTA 270
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 9220 AAGGAGCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 9164
QY 271 TTGCTCAGGTCAAAACAGATGAGCGGCGGATGAGCGCGCGCGCGCGCGCGCG 330
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 9163 GTATATTAAGTTAAAGACCACTGATGATGATGATGATGATGATGATGATG 9104
QY 331 TTGTCACGTTTGTGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 390
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 9103 TTATCTACTATCTTATGAGAAATGAAGAAATTAACAAAGCTTGTATATAGTA 9044
QY 391 GGCACACGTCATTCCTACGAGACCGTGTGACGCGCGCGCGCGCGCGCGCGCGCG 450
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Db 9043 AAGATATTAGTATGATATGAGACTGTGAGTTACAGACGAGTCTTACATTGTA 8984
Oy 451 GCCCGGATAGAGAGAGTCCCGGTGACCTCCGCCCGGCTTACACCTGATG 510
Db 8983 TCACCAATGATAGAGTACAGAGAAAGTATAGCTCAAGTTGGCGCAAGTTCTACAA 8924
Oy 511 CGAACCAAGAGGCGCGGTGTGATGAGGGGTGCGCGCGGTGACCGCGCGAG 570
Db 8923 AACTTAATAGTGTATGAGATTTCTACTAGTGTGTCCAGAGTACTTAAAGGTAA 8864
Oy 571 GTCTGTGTATGAGCGCGCGGACCGCGGTCAACAGCAGCGCGCATGCGCAAGGCGATG 630
Db 8863 GTACTATATATGCGGTGTGAGTCAAGAGAGAACAAATGAGCTAAATTCATAGAGACTA 8804
Oy 631 GCGCGACCGGTAGAGTTTACATCAATCAATGCAAACTTCCGCACTGACCGCGAG 690
Db 8803 GGTGAGATGTTACGATTTTATGATTTTATCCAAAGCGTTTCAACAAATTAAGATGTTTA 8744
Oy 691 TTCTGCGCGCGGATCCACACTGCTACTATCGGCTTACAGCTGAGAGGTGCGGTCAAA 750
Db 8743 TTGCGGAGACGTATACATTAATTAATTCGAAATCCGTTGAATATGATGTTATGATAA 8684
Oy 751 CGTCCGACCTGTATTTGGGCGGTCCGTGCGGAGCGCGCAAGCGCAATTAATGTC 810
Db 8683 CAAAGTATTTATGATATGATGCTAGCTTTTAAATTCAGGCTTAAAGCGCAAGCTTGT 8624
Oy 811 TCGAATTCATCTGTGCGGATATGAACAGGTGCGGTGCTGATGATATACCATCGAC 870
Db 8623 ACAGAGACATATTAACAAATGAAGAAATGAGTCAAGTATTAATGATCTTATGAT 8564
Oy 871 CAGGCGCGGTGTTGCAAGGCTC---ACGACGACACCTTACAGACACCGACGTTGCGC 927
Db 8563 CAGCGCGATTTTGAACCACTGATAAAATTAACAGACATGATGATCTTATATAT 8504
Oy 928 GTGACAGACAGCTTTTACTGCTGTGCGAGATGCCCGCTGCTGCGGAGAGAGCTG 987
Db 8503 AAGCATGTGTGATCATTAATGCAATATGCAATATGCGAGGTGAGTACCGGTACTGCG 8444
Oy 988 ACCTAGCGGCTGACCAAGCAGATGCGGTATGCTGCTGAGCTTGCAGCATGCTGAG 1047
Db 8443 AGCTTAGCTTAAATTAATGCTACGCTTACCTTATGCGCTCATCTGATTAATTAAGGTA 8384
Oy 1048 CGGCGCGCGTCCGCTGCAATCGGCACTAGCCAAAGTCTTTCAGACGACGAGG 1103
Db 8383 AGAGAAGCATTTAAATCAATCAATCAATATGATTAATTAATCAATCAAGG 8328

RESULT      8
ID 167971 standard; DNM: 1074 BP.
AC 167971;
DT 15-JUL-1997 (first entry)
DE H. pylori membrane protein ORF 05cp20518orf1.
KW Vaccine; prevention; treatment; infection; identification;
   binding compound; bacterium; life cycle; activator; bacteris;
   inhibitor; duodenal ulcer disease; life cycle; activator; bacteris;
   membrane; amino acid; metabolism; ds.
OS Helicobacter pylori.
FH Key
FT 1. 1074
FT cds
FT /*tag= a
FT /*note= "no stop codon given"

MO9640893-A1.
PD 19-DEC-1996.
PF 06-JUN-1996; U09122.
PR 07-JUN-1995; US-487032.
PR 01-APR-1996; US-630405.
PA (ASTR ) ASTRA AB.
PI Berglindh OT, Smith D, Wellgaard B.;
DR WPI: 97-052306/05.
DR P-BSDB: W20718.
PT Helicobacter pylori nucleic acid sequences and related
   polypeptide(s) - useful for vaccines to treat or prevent H. pylori

```

```

PT Infection, and to detect Helicobacter
PS Claim 27; Page 819; 1481p; English.
CC The present sequence encodes a Helicobacter pylori membrane
CC protein likely to contain four membrane spanning regions.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds,
CC useful as potential H. pylori life cycle activators or inhibitors.
CC The genomic sequence of H. pylori (ATCC 35679) was determined from
CC overlapping contigs generated by mechanically shearing the
CC bacterial DNA. The sequences were analysed for ORF of at least 180
CC nucleotides, and the predicted coding regions defined by computer
CC evaluation. To identify likely H. pylori antigens for vaccine
CC development, the amino acid sequences predicted from various ORF
CC were analysed for significant homology to other known or exported
CC membrane proteins. Having identified and determined the sequences
CC of interest, particular regions can be isolated from H. pylori by
CC PCR amplification for recombinant polypeptide production, e.g. in
CC E. coli hosts.
SQ Sequence 1074 BP; 302 A; 189 C; 297 G; 286 T;

```

```

Query Match      7.7%; Score 94.8; DB 1; Length 1074;
Best Local Similarity 46.1%; Pred. No. 1,2e-11;
Matches 370; Conservative 0; Mismatches 417; Indels 15; Gaps 1;

```

```

Oy 265 GATTATTTGCTCAAGGTCAAGAACCGATAGCGGCGAATACGGCGCTCGACACGG 324
Db 133 GATTGTTGCTCAAGGTCAAGAACCGCTTAAAGCATGATACCTTGTCTCAAGAAAA 192
Oy 325 CAGATCTTGTTCACGCTTCTTCATTTGCGCGCGTACGCTGTCACAGCATGCTGTTG 384
Db 193 GCGACTGTGTTAGTATTGATTTAGCTATCAAAAACCTTGTGAAATTTTAT 252
Oy 385 GATTCGCGACACGCTCAATTCCTACGACACGCTCCAGACCGCGGACGCTACCC 444
Db 253 AATAAAAAATCACTTCTTATTTGCACTGAAACCAATTCGCGGCGCTTAAAGACTACCT 312
Oy 445 CTGCTTGCCTGATGAGGAGAGTGCAGCGTGCAGTCCGCGCGCGCGAGTGGG 496
Db 313 ATTTAGGCGCTATACCTGTGTGCTGAGAGGTTGCTGCGCATTTAGTACACATTAT 372
Oy 496 -----GCTTACCACTGATGCGAAACCAAGGGCGCGGTGTGCTGATGGCGGGTG 549
Db 373 TTACTGCTTTTAAAGCATGTTAAAGGTTTATGGTAAAGGGGCTCATGAGGGTTTG 432
Oy 550 CCGGCGTGCAGACCGCGCGACCTCTGTGATGCGCGCGCGCGCGCGCTTACAGCA 609
Db 433 TCGGCTGCGCAAGGGCTTAAATGCTGTAATTCGAGGCGGTGTGTTGGATGAGAGC 492
Oy 610 GCCCGCATGCGCAAGCGCATGCGCGCGACCTTACAGGTTCTAGACATCAATGACAA 669
Db 493 GCGAAAGCTTAAAGCAATGAGGCGCTTAAAGATTTTAAAGATTTAGACTAGCTTAA 552
Oy 670 CTTGCGCACTGACGCGGATCTGCGGCGGATCCACACTGCTACTATCGGCTTAC 729
Db 553 TTACAAACCAACCTTATTAATCAATTTATGATTTAGAAAGCTTAAAGGTGAAGAGCC 612
Oy 730 GAGCTGAGGGTGCCTCAACGCTGCGAGCTGAGTATGGGCGGCTGTCGCGAGG 789
Db 613 AATATCATTAAGGCTTAAAGGCGGCTGAGGAGGAGGAGGCTGCTGTATACAGG 672
Oy 790 GCGAAGGACCAATTTAGTCTGCAATTCATTTGCGGCAATTAAGAACGAGTGGCGTA 849
Db 673 AGCCAAACCTTAAAGTATCTTAAAGAGGATTTAAATTAATGACAGACCAAGAGGTA 732
Oy 850 CTGTGATATAGCATGACAGGCGGCTGTTTCGAGGCTCAGCAGCAGCAGCTTAC 909
Db 733 GTCAATGATGTGTGAGTTTAAAGGAGGAGGAGTAAAGGCAATTCAGCAAGAGCAT 792
Oy 910 GACACCGGAGTGTGCGCTGACAGACGCTGTTTACTGCTGAGGAGACATCCCGCC 969
Db 793 TCTAACCGGATGTGTGAGAAAGTTTGTTCATATGAGCTGCGCAAGATGCGAGG 852

```


KW hybridisation; detection; vaccine; immunisation; infection; ss.
 OS Mycobacterium sp.
 PD MO9909186-A2.
 PD 25-FEB-1999.
 PF 14-AUG-1998; F01813.
 PR 11-SEP-1997; FR-011325.
 PR 14-AUG-1997; FR-010404.
 PA (INSP) INST PASTEUR.
 PI Giequel B, Llm EM, Pellicic V, Portnoi D, Goguet de la Salmoniere Y,
 PI Guigueno A;
 DR MPI: 99-181045/15.
 DR P-PSDB: Y04952.
 PT Mycobacterial DNA vectors containing reporter constructs - for
 PT identifying coding or promoter sequences involved in
 PT infection-associated protein expression
 PS Claim 22; Fig 41S; 309pp; French.
 CC Sequences X34001-X34252 represent nucleic acids encoding secreted
 CC proteins from various Mycobacterium species microorganisms. The
 CC nucleotide sequences can be used as primers and probes for methods
 CC for detecting and identifying mycobacteria, especially belonging to
 CC the M. tuberculosis complex. The encoded proteins can be used in
 CC vaccines for immunisation against a bacterial or viral infection.
 SQ Sequence 1413 BP; 187 A; 459 C; 540 G; 227 T;

Query Match 4.1%; Score 50.4; DB 1; Length 1413;
 Best Local Similarity 43.0%; Pred. No. 0.021;
 Matches 246; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

QY 69 CGGATTCGACGAGACCAAAACAAAGATTTCGGGTCATCCAGCCGCGCGGCT 128
 DB 69 CGGATTCGACGAGACCAAAACAAAGATTTCGGGTCATCCAGCCGCGCGGCT 128
 QY 129 CGGGAATTAACCCGCTGTCGATGAGTGTCTATCCAGCCGCGCGGAGGCTC 188
 DB 129 GCGCGGCTGTCGATGAGTGTCTATCCAGCCGCGCGGAGGCTC 188
 QY 189 GCGTATCAACGAGCGGATTTCAAGCGCGGAGCGGAGCTGTCGCGGAGCA 248
 DB 189 GCGTATCAACGAGCGGATTTCAAGCGCGGAGCGGAGCTGTCGCGGAGCA 248
 QY 249 GGTGTGGCGGACGATTTATGCTCAAGGTCAAGAACGATGCGGAGGATACGG 308
 DB 249 GGTGTGGCGGACGATTTATGCTCAAGGTCAAGAACGATGCGGAGGATACGG 308
 QY 309 CGGCTGCGACACGGGAGATTTGTCACGTTCTTGATTTGGCGGTCACGCTTG 368
 DB 309 CGGCTGCGACACGGGAGATTTGTCACGTTCTTGATTTGGCGGTCACGCTTG 368
 QY 369 CACGATGCGTGTGATTCGCGACGACGATTCGCTGAGAGCGGAGCGGAGCG 428
 DB 369 CACGATGCGTGTGATTCGCGACGACGATTCGCTGAGAGCGGAGCGGAGCG 428
 QY 429 CGAGCGGACATACCCCTGCTTCCCGATGAGCGAGTCCGCGGTCGACGCGGCA 488
 DB 429 CGAGCGGACATACCCCTGCTTCCCGATGAGCGAGTCCGCGGTCGACGCGGCA 488
 QY 489 GGTGGCGGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 548
 DB 489 GGTGGCGGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 548
 QY 549 GCGCGGCTGAGACCGCGGCGGAGCTGTGATGCGCGCGGAGCGGAGCGGAGCG 608
 DB 549 GCGCGGCTGAGACCGCGGCGGAGCTGTGATGCGCGCGGAGCGGAGCGGAGCG 608
 QY 609 AGCGCGCATCGGACGCGGATGCGGCGGAGCG 640
 DB 609 AGCGCGCATCGGACGCGGATGCGGCGGAGCG 640

RESULT 14
 X34206
 ID X34206 standard; DNA: 1722 BP.

AC X34206:
 DE 06-JUL-1999 (first entry)
 DT Mycobacterium species nucleic acid sequence 41T.
 KW Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
 KW hybridisation; detection; vaccine; immunisation; infection; ss.
 OS Mycobacterium sp.
 PD MO9909186-A2.
 PD 25-FEB-1999.
 PR 14-AUG-1998; F01813.
 PR 11-SEP-1997; FR-011325.
 PR 14-AUG-1997; FR-010404.
 PA (INSP) INST PASTEUR.
 PI Giequel B, Llm EM, Pellicic V, Portnoi D, Goguet de la Salmoniere Y,
 PI Guigueno A;
 DR MPI: 99-181045/15.
 DR P-PSDB: Y04953.
 PT Mycobacterial DNA vectors containing reporter constructs - for
 PT identifying coding or promoter sequences involved in
 PT infection-associated protein expression
 PS Claim 22; Fig 41T; 309pp; French.
 CC Sequences X34001-X34252 represent nucleic acids encoding secreted
 CC proteins from various Mycobacterium species microorganisms. The
 CC nucleotide sequences can be used as primers and probes for methods
 CC for detecting and identifying mycobacteria, especially belonging to
 CC the M. tuberculosis complex. The encoded proteins can be used in
 CC vaccines for immunisation against a bacterial or viral infection.
 SQ Sequence 1722 BP; 231 A; 548 C; 657 G; 286 T;

Query Match 4.1%; Score 50.4; DB 1; Length 1722;
 Best Local Similarity 43.0%; Pred. No. 0.022;
 Matches 246; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

QY 69 CGGATTCGACGAGACCAAAACAAAGATTTCGGGTCATCCAGCCGCGGCGT 128
 DB 69 CGGATTCGACGAGACCAAAACAAAGATTTCGGGTCATCCAGCCGCGGCGT 128
 QY 378 GCGGCTGCGGCGGCGGCTGTCGATGAGTGTCTATCCAGCCGCGGAGGCTC 437
 DB 378 GCGGCTGCGGCGGCGGCTGTCGATGAGTGTCTATCCAGCCGCGGAGGCTC 437
 QY 129 GCGGGAATTAACCCGCTGTCGATGAGTGTCTATCCAGCCGCGGAGGCTC 188
 DB 129 GCGGGAATTAACCCGCTGTCGATGAGTGTCTATCCAGCCGCGGAGGCTC 188
 QY 438 GCGGCTGTCGATGAGTGTCTATCCAGCCGCGGAGGCTC 497
 DB 438 GCGGCTGTCGATGAGTGTCTATCCAGCCGCGGAGGCTC 497
 QY 189 GGTGTGGCGGACGATTTATGCTCAAGGTCAAGAACGATGCGGAGGATACGG 248
 DB 189 GGTGTGGCGGACGATTTATGCTCAAGGTCAAGAACGATGCGGAGGATACGG 248
 QY 498 GGTGTGGCGGACGATTTATGCTCAAGGTCAAGAACGATGCGGAGGATACGG 557
 DB 498 GGTGTGGCGGACGATTTATGCTCAAGGTCAAGAACGATGCGGAGGATACGG 557
 QY 249 GGTGTGGCGGACGATTTATGCTCAAGGTCAAGAACGATGCGGAGGATACGG 308
 DB 249 GGTGTGGCGGACGATTTATGCTCAAGGTCAAGAACGATGCGGAGGATACGG 308
 QY 558 GGTGTGGCGGACGATTTATGCTCAAGGTCAAGAACGATGCGGAGGATACGG 617
 DB 558 GGTGTGGCGGACGATTTATGCTCAAGGTCAAGAACGATGCGGAGGATACGG 617
 QY 309 CCGGCTGCGACACGGGAGATTTGTCACGTTCTTGATTTGGCGGTCACGCTTG 368
 DB 309 CCGGCTGCGACACGGGAGATTTGTCACGTTCTTGATTTGGCGGTCACGCTTG 368
 QY 618 CCGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 677
 DB 618 CCGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 677
 QY 369 CACGATGCGTGTGATTCGCGACGACGATTCGCTGAGAGCGGAGCGGAGCG 428
 DB 369 CACGATGCGTGTGATTCGCGACGACGATTCGCTGAGAGCGGAGCGGAGCG 428
 QY 678 GGTGTGGCGGACGATTTATGCTCAAGGTCAAGAACGATGCGGAGGATACGG 737
 DB 678 GGTGTGGCGGACGATTTATGCTCAAGGTCAAGAACGATGCGGAGGATACGG 737
 QY 429 CGAGCGGACATACCCCTGCTTCCCGATGAGCGAGTCCGCGGTCGACGCGGCA 488
 DB 429 CGAGCGGACATACCCCTGCTTCCCGATGAGCGAGTCCGCGGTCGACGCGGCA 488
 QY 738 CTACGAAACCTTCGCGCGGCGGAGCTGTGATGCGCGCGGAGCGGAGCGGAG 797
 DB 738 CTACGAAACCTTCGCGCGGCGGAGCTGTGATGCGCGCGGAGCGGAGCGGAG 797
 QY 489 GGTGGCGGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 548
 DB 489 GGTGGCGGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 548
 QY 798 ATTGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 857
 DB 798 ATTGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 857
 QY 549 GCGCGGCTGAGACCGCGGCGGAGCTGTGATGCGCGCGGAGCGGAGCGGAGCG 608
 DB 549 GCGCGGCTGAGACCGCGGCGGAGCTGTGATGCGCGCGGAGCGGAGCGGAGCG 608
 QY 858 TCGACTCAAGGAATTTGGGTGATGCGGATGCGCGCTGCTCGCGGAGAGTACTTC 917
 DB 858 TCGACTCAAGGAATTTGGGTGATGCGGATGCGCGCTGCTCGCGGAGAGTACTTC 917
 QY 609 AGCGCGCATCGGACGCGGATGCGGCGGAGCG 640
 DB 609 AGCGCGCATCGGACGCGGATGCGGCGGAGCG 640
 QY 918 GGTACAGTCCACACGACGACGCGGCTGCGG 949
 DB 918 GGTACAGTCCACACGACGACGCGGCTGCGG 949

Fri Jun 23 09:31:28 2000

us-09-362-485-3.rn1

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: June 22, 2000, 15:01:15 ; Search time 341.15 Seconds
(without alignments)
470.560 Million cell updates/sec

Title: US-09-362-485-3
Sequence: 1 ATCTTCAGATTAATCGAAC.....ANGAGAGCATGATGTCGNC 1235

Scoring table:
IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/5C.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/5D.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/6.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/PCUTUS.COMB.seq:*
7: /cgn2_6/ptodata/2/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47.6	3.9	1620	US-08-461-775-10	Sequence 10, Appl
2	47.6	3.9	28958	US-08-258-261B-6	Sequence 6, Appl
3	47.6	3.9	28958	US-08-456-837-6	Sequence 6, Appl
4	47.6	3.9	28958	US-08-457-342-6	Sequence 6, Appl
5	47.6	3.9	28958	US-08-457-646A-6	Sequence 6, Appl
6	47.6	3.9	28958	US-08-458-076A-6	Sequence 6, Appl
7	47.6	3.9	28958	US-08-764-233A-4	Sequence 4, Appl
8	47.6	3.9	28958	US-08-457-335A-6	Sequence 6, Appl
9	47.6	3.9	28958	US-08-729-214-6	Sequence 6, Appl
10	47.6	3.9	49377	US-08-764-233A-1	Sequence 1, Appl
11	46.6	3.7	2582	US-08-816-105A-2	Sequence 2, Appl
12	45.4	3.7	833	US-08-403-852D-3	Sequence 3, Appl
13	45.4	3.7	5392	US-08-403-852D-1	Sequence 1, Appl
14	42.4	3.4	1832	US-08-403-852D-6	Sequence 6, Appl
15	42.4	3.4	1832	US-08-173-508-3	Sequence 3, Appl
16	42.4	3.4	2185	US-08-265-310-3	Sequence 3, Appl
17	41.2	3.3	2185	US-08-461-775-11	Sequence 11, Appl
18	41.2	3.3	2668	US-07-642-734C-3	Sequence 3, Appl
19	41.2	3.3	20235	US-07-642-734C-1	Sequence 1, Appl
20	40.8	3.3	44377	US-08-804-227C-7	Sequence 7, Appl
21	40.8	3.3	44377	US-08-804-198-1	Sequence 1, Appl
22	40.6	3.3	4257	US-08-690-473-1	Sequence 1, Appl
23	40.6	3.3	12001	US-08-458-568A-11	Sequence 11, Appl
24	40.2	3.3	2064	US-08-343-428-1	Sequence 1, Appl
25	40.2	3.3	11219	US-07-642-734C-1	Sequence 1, Appl
26	39.6	3.2	474	US-08-403-852D-14	Sequence 14, Appl
27	39.6	3.2	1524	US-08-461-775-10	Patent No. 5512669

28	39.6	3.2	1525	7	5229279-1	Patent No. 5229279
29	39.4	3.2	459	7	US-08-387-942C-35	Sequence 35, Appl
30	39.4	3.2	1998	7	5212296-8	Patent No. 5212296
31	39.4	3.2	12588	4	US-08-387-942C-1	Sequence 1, Appl
32	39.2	3.2	420	1	US-08-470-179-148	Sequence 148, App
33	39.2	3.2	2353	6	PCT-US92-06840-1	Sequence 1, Appl
34	39.2	3.2	43280	3	US-08-804-227C-1	Sequence 1, Appl
35	38.8	3.1	8051	4	US-08-576-656A-2	Sequence 2, Appl
36	38.8	3.1	11219	2	US-07-642-734C-1	Sequence 1, Appl
37	38.6	3.1	1187	2	US-08-440-856A-2	Sequence 2, Appl
38	38.6	3.1	3231	2	US-08-074-121-4	Sequence 4, Appl
39	38.6	3.1	3231	6	PCT-US94-06447-4	Sequence 4, Appl
40	38.4	3.1	1215	4	US-08-947-726A-1	Sequence 1, Appl
41	38.4	3.1	2048	5	US-08-776-251-1	Sequence 1, Appl
42	38.2	3.1	2109	4	US-08-555-568B-20	Sequence 20, Appl
43	38	3.1	30001	1	US-08-125-468-1	Sequence 1, Appl
44	38	3.1	30001	3	US-08-474-933-1	Sequence 1, Appl
45	37.8	3.1	2588	3	US-08-796-414B-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-461-775-10
Sequence 10, Application US/08461775
Patent No. 5858773
GENERAL INFORMATION:
APPLICANT: MAZODIER, Philippe
APPLICANT: GUGLIEMTI, Gerard
TITLE OF INVENTION: REGULATORY NUCLEOTIDE SEQUENCE OF THE
TITLE OF INVENTION: INITIATION OF TRANSCRIPTION
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22131-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,775
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/050,313
FILING DATE: 10-MAY-1993
APPLICATION NUMBER: FR 9011186
FILING DATE: 10-SEP-1990
ATTORNEY/AGENT INFORMATION:
NAME: Crane Feunty, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 010830-035
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1620 base pairs
TYPE: nucleic acid
STRANDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1620
US-08-461-775-10

Query Match	3.9%	Score 47.6;	DB 3;	Length 1620;
Best Local Similarity	45.7%	Pred. No. 0.0046;		
Matches 321; Conservative	0;	Mismatches 369;	Indels 12;	Gaps 4

QY	114	CACCCGCGCGGGTGGCGGAACTCAACCGCTGCTGGCATTGAGGTGCTCATCCAGCAGG	173
Db	879	CACCTTACCGGTGTCCACCTCATTCGCCAGGAGGTGGGCTTCAAGCTGGACCAAGCCGG	938
QY	174	TGCCGAGAGGGGCTCGCTTATCAACCGAGCGGATTTTCAAGGGCGCAGGGCCCAACTGT	233
Db	939	TCTGAGAGTGTGTGGGACACCGCCCGCGGCTACCGTCAACCAAGGAGACAGACCATCTGT	998
QY	234	CGGCACCGCGGACCAAGTGTGGGCCGACGCTGATTTATTGTCTCAAGTCAAGAACCAT	293
Db	999	GGAGGGGGGGGCAACACGAGGACGTCCAGGGCCCGCGCTGGCCAGATCAAGCCGAAT	1058
QY	294	AGCGCGGAATACGGCGCGCTGGGACACGGGGAGATCTTGTTACAGTTTCTTCATTGGC	353
Db	1059	CGAATCCACGACCTCGGACCTGGGACCGCGAGAAAGCTCCAGAGACCGCTTGGCAACTGGC	1118
QY	354	C-GCGTCACGTGCTTGCCACCGATGCTGTTGGATTCCGGACACAGTCATTTGCTACG	412
Db	1119	CGGGGGGTGTGCTGTATCTTCCGCTCGCGCGGGCCACCGAGTCTGAGCTGAAGAGCCGA	1178
QY	413	AGACCGTCCAGAGCGCGGACGGGGCACTACCCCTGTCGCCGAGTAGACGAA--GTGCG	470
Db	1179	GCACCGTCTGAGAGAACCCCATCTTCGCGAGACCCCGCGCGGGTGTGAGAGGGCATGTCTC	1238
QY	471	CGGTGACTCGCGGCCACGATTGGCGTTACCA--CTGATGCGAACCACAGGGGGCG	527
Db	1239	CGGTGTGGGTCTCCCGCTGTCTCCACGCGCTCAAGGTCTGTGAGACAACTCGCGCCGAC	1298
QY	528	CGGTGTCTGATGGGCGGGGTGCGCGGCGTCGAACCGCGGAGCGTGGGTGATGG--	585
Db	1299	CGGGAGAGAGGCCACCGGTGTCCGGGTCTGCGCCGCCGCCGCTGTGAGCCGCTCGCTG	1358
QY	585	--GCGCGGACCGCGCGCTACAAACGACCGCCATCGCCAAAGGCGATGGCGCGACCGT	641
Db	1359	GATGCGCGAAGAACCGCGCTCTCAAGGGGTACGTATTCATACCAAGGTGGCGAGCTGA	1418
QY	642	TACGGTTCTGACATCAACATGCAAAACTTGGGCAACTCGACGGCCGAGTTCTG-GGGCG	701
Db	1419	CAAGGGCCAGGGCTTCAACCGGCGCAACCGCGAGTACGGCGACCTGTGTCAAGCCGGGT	1478
QY	702	GATCCACTCTGCTACTCATCGGCGCTCAAGGAGCTCGAGGGTGGCGTCAACGCTGCGACCT	761
Db	1479	CATCGACCCCGGTCAAGGTCAACCGGTGTCCGCCCTCGAAGAACGGGGCTCATAGCTCTCT	1538
QY	762	GGTGATTGGGGCGGTCTGTGTGCGAGGCGCAAGGACCCAA	803
Db	1539	GTCTCTGACACCGAGACCTGTGTCTGTGGAAGACGGGGCGGA	1580

US-08-258-261B-6
2
Sequence 6, Application US/08258261B
Patent No. 5639949
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Galfney, Thomas Deane
APPLICANT: Lam, Stephen Ting
APPLICANT: Hammer, Phillip E.
APPLICANT: Uknes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
antipathogenic substances
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Gelgy Corporation

STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Vers
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/258,261B
FILING DATE: 08-JUN-1994
CLASSIFICATION: 800

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995

ATTORNEY/AGENT INFORMATION:
NAME: Elmer James Scott
REGISTRATION NUMBER: 36,129

REFERENCE/DOCKET NUMBER: CGC 1506/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO

OS-08-258-261B-6

Query Match	3.98;	Score 47.6;	DB 1;	Length 28958;
Best Local Similarity	44.1;	Pred. No. 0.01;		
Matches 286;	Conservative 0;	Mismatches 359;	Indels 3;	Gaps 2;

OY	148	GGATAGGTGGTCTATCCAGGACAGTGCAGAGGGGCTGGCTATCACCAGCGGGGAT	207
Db	25366	GCCACGACGACCTCAACCTCGCGGATGTGGCTTTGGCTGGCCACACCGCGGCCAC	253655
OY	208	TTCAAGCGGAGCGCGCGACCTGGTCGGACCGCGCCAGCGAGGTGTGGCGCAGCTGAT	267
Db	25366	TTGAGGACCGCGCGCTCTCTGTAGCCCAAACTCCGACGAGTCTCTCTCGCGCTCGAC	254255
OY	268	TTATGTGTAAGTTCAAAACACCATACCGCGGAAATTCGGCGCGCTCTCGACACGGGCGAG	327
Db	25426	TGCTTCCCCACGACACAGCGCGCGCCCCGAGACCGTCTCGGACGAGCGGAAACCCAGCGG	254855
OY	328	ATCTTTGTCAGTTCCTTGCATTGTGGCCCGACAGCTGCTTCACCGATGGGTGTGGAT	387
Db	25486	AAGCTGCTTGTGCTTCTTCTGGGCAAGGTGTGGAGTGGGAAGGATGGCCCTCTGGCG	255455
OY	388	TCCGGCACACACGTCATTCCTCTACGACACGTCACAGACCGCCGACGGCGCATACCCCTG	447
Db	25546	CTGACATCTCTGCGCCCTCTTCGCGCCTCAGTCTGAAAGATCTCGAGCCCGCGCTCGCTCT	256055
OY	448	CTTACCCCGATGAGCAATGTCGCGCGGTGACATCGCGCGCCAGAGTTGGCGGTTACAC--C	505
Db	25606	CAGTGTAGTGGAGGCTGCTGCTGCCCTCTCGGCGCGACAGAGGGCGCCCTCTCCCTCGAC	256655
OY	506	TGATCGAACCCAAAGGGGGCGCGGTGTGCTGATGGCGGGGTGCCCGCGCTGGAACCGG	565
Db	25666	CGGTCACACCTCTACAGCGCCCGCTTTCCGTCATGTGCTCTCCCTGGCGGCGCTCTGG	257255
OY	566	CCGACG--TCGTGGTGAATCGCGCGCGGACCGCGCGGATCAACGCAAGCCCATCGCCACAC	624
Db	25726	CGCTCGCTGGGCTAGAGCCCGCGCGCTGTGGCCACACTCAGGGCGAGATGGCGCGCC	257855
OY	625	GGCATGGGCGCCAGCGTTACGGTTTATAGATCAACATCGACAACTTCGGCAACTCGAC	684

Db 25786 GCCTTGCTGCAGACGCTCTCTCCCTGAGAGCGCGCCGCAATCGCCCTGCGCAGC 25845
QY 685 GCCGAGTTCTGCGCGCGGATCCACACTGCTACTCATCGGCTAGAGCTCGAGGGTGC 744
Db 25846 AAGCGCTACACACCGTCCGCGGAGGAGGCGCATGGCCGCTCCAGCTCGCGCTCC 25905
QY 745 GTCAAAAGTCCGACCTGTGATTTGGGCGCTCTGCTGCGAGCGCC 792
Db 25906 GACCTCCAGACTACCTGCTCCCTGGGCGAGCAGGCTCTCATCGCC 25953

RESULT 3

US-08-456-837-6
; Sequence 6, Application US/08456837
; Patent No. 5643774
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,837
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/457,205
; FILING DATE: 01-JUN-1995
; APPLICATION NUMBER: 08/258,261
; FILING DATE: 08-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Rimer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-456-837-6

Query Match 3.9%; Score 47.6; DB 1; Length 28958;
Best Local Similarity 44.1%; Pred. No. 0.01;
Matches 286; Conservative 0; Mismatches 37; Indels 3; Gaps 2;

QY 148 GGCCATGAGTGTCTATCC/ GCAGGTGCGGAGAGAGGCTCGGTATACGACGCGGAT 207
Db 25306 GCCACGAGAGACTACACCTCCGCGATGTGGCTATTCGCTGACACACCGCGCCAC 25385
QY 208 TTCAGGCGGAGAGCGCGCACTGTGTGGCAGCGCCGACAGAGTGTGGCGCAGCTGAT 267
Db 25366 TTCAGGCGGAGAGCGCGCTGTGTAGCCCAACCGCGAGAGACTCTCTCGCGGCTGAC 25425
QY 268 TTATGCTCAGAGTCAAAACAGATAGGCGGGAATAGCGCGCTGTGACAGGCGAG 327
Db 25426 TCGCTGCGGAGAGAGAGCGCGCGCGGAGACAGCTGCTGAGAGGAGGAGGAGCGC 25485
QY 328 ATCTTGTCAGCTTTCTTGATTTGGCGCGGTCAGCTGCTTGACGAGATGCTTGGAT 387
Db 25486 AAGCTGCTTCTGCTGCTTCTGCGGAGAGCTGCGAGTGGAGAGGATGGCCCTCTGCTG 25545
QY 388 TCCGCGACAGCTCATTTG/ TACGAGACCTTCAGACCGCGCGAGCGGCGACTACCCCTG 447
Db 25546 CTCGACTCTCTGCGCGCTCTTCCGCGCTCAGCTGGAAGCATGCGAGCGCGCTGCTCT 25605
QY 448 CTTGCCCGATGAGGAGAGTCCGCGGTGACTCGCGCGGCGGAGTTGGGCTTACGAC--C 505
Db 25606 CACGTGAGATGAGAGCTGTGCTGCGCTCTGCGCGCGGAGAGGCGCGCTCTCTGAC 25665
QY 506 TGTGCGAACCAGAGGCGCGGTGTGCTGATGGGCGGCTGCGCGCTGCAACCGG 565
Db 25666 CCGGTGAGCTGTGACAGCGCGCTCTTGGCGGTGATGTCTCTGCGCGCGCTCTG 25725
QY 566 CCGAGC-TGTGTGATGGCGCGCGGACACCGCGGCTCAACGAGCGCGGATGCCAAC 624
Db 25726 CCTGCTGCTGCGGTGAGAGCGCGCGCTGTGCGCGGACAGTCAAGGAGATGCGCGC 25785
QY 625 GGCATGGGCGGCGAGCTTACGTTCTAGACATCAACATGACAACTTCGCAACTGAC 684
Db 25786 GCCTTGCTGCGAGGCGCTCTCTCCCTGAGAGAGCGCGCGCGCTCTGCGCAGC 25845
QY 685 GCCGAGTTCTGCGCGCGGATCCACACTGCTACTCATGCGCTAGAGCTGAGGTGCC 744
Db 25846 AAGCGCTACACACCGTCCGCGGAGGCGCATGGCGCGCTGACAGCTGCGCGCTCC 25905
QY 745 GTCAAGCGTCCGACCTGTGATTTGGGCGCTCTGCTGCGAGCGGCGGCGGCGGCG 792
Db 25906 GACCTCCAGACTACCTGCTCCCTGGGCGAGCAGGCTCTCATCGCC 25953

RESULT 4

US-08-457-342-6
; Sequence 6, Application US/08457342
; Patent No. 5652898
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA: 25906
APPLICATION NUMBER: US/08/457,342
FILING DATE: 01-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-457-342-6

Query Match 3.9%; Score 47.6; DB 1: Length 28958;
Best Local Similarity 44.1%; Pred. No. 0.01;
Matches 286; Conservative 0; Mismatches 359; Indels 3; Gaps 2;
QY 148 GCCCATGAGTCTCTATCCAGGAGTGGCGGAGGCGTCCGCTATCCAGCGCGGAT 207
DB 25306 GCCCAGAGAGCTCACTCCGCGATGCGCTATCGCTGCGCCACCGCGCGCCAC 25365
QY 208 TTCAGGCGGCGAGCGCGCAATGTCGCGACCGCGCGACGAGTGTGGCGCGACGCTGAT 267
DB 25366 TTCGAGCAGCGCGCGCTCTCGTACGCCCAACCGGAGAGAGCTCTCGCGCGCTGAGC 25425
QY 268 TTATGCTCAGGTCAAGAACCGATAGCGGCGGAATACGCGCGCTCGGACAGCGGCGAG 327
DB 25426 TCGCTCGCCAGCAAGCGCGCGCGCGCGAGCGCTCTCGAGCGAGCGAAGCCACGCGC 25485
QY 328 ATCTGTACGCTCTGCTGATTTGGCGGCGTCAAGCGCTTGCACGATGCTGTGAT 387
DB 25486 AAGCTGCTGCTCTTCTCTGGGCAAGGCTGCAAGTGGAAAGGATGCGCTCTGCTG 25545
QY 388 TCCGAGCAGCTCAATGCTACGAGACCGTCCAGACCGCGCGCGACGCTACCGCTG 447
DB 25546 CTCGACTCTCGCGCGCTCTTCGCGCTCAGCTCGAAGCATGGAAGGCGCGCTCGCTCT 25605
QY 448 CTGGCCCGATGAGAGAGTCCGCGCTGCACTGCGCGCGCGCGCGCGCTTACCGAC--C 505
DB 25606 CAGCTGAGTGAAGCTCTCGCGCTCTGCGCGCGCGAGAGAGGCGCGCGCTCTCGAC 25665
QY 506 TGATGGAACCAAGGCGCGCGGCTGCTGATGGCGGCGCGCGCGCGCGCGCGCGCGG 565
DB 25666 CGGCTGAGCTGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 25725
QY 566 CCGAGC--TCGTGATGTCG 624
DB 25726 CGCTGCTGCGGAGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 25785
QY 625 GCGATGCG 684
DB 25786 GCGTCTGCTGAGGCGCGCTCTCTCGCGTGAAGCGCGCGCGCGCGCGCGCGCGCGCG 25845
QY 685 GCGGAGTGTGGGCG 744
DB 25846 AAGCGCTCAACCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 25905

QY 745 GTCAAGCTGCCGACCTGCTGATGCGGCGGCTCTCTGATCCAGCGCGCG 792
DB 25906 GACCTCCAGACTACTGCTCTCCCTGGGCGGAGCGGCTTCATCGCC 25953
RESULT 5
US-08-457-646A-6
Sequence 6, Application US/08457646A
Patent No. 5679560
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Gathney, Thomas Deane
APPLICANT: Lam, Stephen Ting
APPLICANT: Hammer, Phillip E.
APPLICANT: Umes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 93/08/457,646A
FILING DATE: 01-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-457-646A-6
Query Match 3.9%; Score 47.6; DB 1: Length 28958;
Best Local Similarity 44.1%; Pred. No. 0.01;
Matches 286; Conservative 0; Mismatches 359; Indels 3; Gaps 2;
QY 148 GCCCATGAGTCTCTATCCAGGAGTGGCGGAGGCGTCCGCTATCCAGCGCGGAT 207
DB 25306 GCCCAGAGAGCTCACTCCGCGATGCGCTATCGCTGCGCCACCGCGCGCCAC 25365
QY 208 TTCAGGCGGCGAGCGCGCAATGTCGCGACCGCGCGACGAGTGTGGCGCGACGCTGAT 267
DB 25366 TTCGAGCAGCGCGCGCTCTCGTACGCCCAACCGGAGAGAGCTCTCGCGCGCTGAGC 25425


```

: Sequence 4, Application US/08764233A
: Patent No. 5716849
: GENERAL INFORMATION:
: APPLICANT: Ligon, James M.
: APPLICANT: Schupp, Thomas
: APPLICANT: Beck, James J.
: APPLICANT: Hill, Dwight S.
: APPLICANT: Neff, Snezana
: APPLICANT: Ryals, John A.
: TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Ciba-Geigy Corporation
: STREET: 520 White Plains Road, P.O. Box 2005
: CITY: Tarrytown
: STATE: NY
: COUNTRY: USA
: ZIP: 10591
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/764,233A
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/729,214
: FILING DATE: 09-OCT-1996
: APPLICATION NUMBER: US 08/258,261
: FILING DATE: 08-JUN-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Meigs, J. Timothy
: REGISTRATION NUMBER: 38,241
: REFERENCE/DOCKET NUMBER: 1506/CIP6
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (919) 541-8587
: TELEFAX: (919) 541-8689
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 28958 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: ORGANISM: Sorangium cellulosum
: IMMEDIATE SOURCE:
: CLONE: P98/1
: US-08-764-233A-4

```

Query Match 3.9%; Score 47.6; DB 2; Length 28958;
 Best Local Similarity 44.1%; Pred. No. 0.01;
 Matches 286; Conservative 0; Mismatches 359; Indels 3; Gaps 2;

```

QY 148 GCCCATGAGTGTCTATCATCAGCAGGTCGCCGAGAGAGGCTCGCTATCCAGACGGGAT 207
DB 25306 GCCCAGACGACTCACCCTCGCGATGTGGCTTATTCCTGCGCCACACCGCGGCCAC 25345
QY 208 TTCAGAGCGCGAGCGCGCACTGTGTGGCAGCGCCGACAGAGTGTGGCCGACGCTGAT 267
DB 25366 TTCGAGACACCGCGCGCTCTGTAGCCCAACACCGAGAGAGCTCTCTCCGCGTCGAC 25425
QY 268 TTATGCTCAAGGTCAAGAAACGATAGCGGGGGAATAGCGCGCTCTGGAGACGGGACG 327
DB 25426 TCGCTGCGCCAGACAGCGCGCGCGAGACCGCTCTCTGAGGAGCGGAGCGACGAC 25485
QY 328 ATCTGTACGCTCTTGTGATTTGGCCGCGCTGCTTGCACCGATGCGTTGTGAT 387
DB 25486 AAGCTGCTGTCTTCTTCTTCTGCGGACGAGCTCTGAGTGGAGAGGATGCGCTCTGCTG 25545

```

```

QY 388 TCCGACACCGATCAATTGCTTACAGAGACGCTCCAGACCGCGGAGCTACCCCTG 447
DB 25546 CTCGACTCTTCCCGCTTTCGGGCTTACGCTCGAAGATCGAGCGGCTGCTCT 25605
QY 448 CTTGCCCGATGAGGAGGAGCGGCTGACATCGCGCCGAGGTTGGGCTTACAC -C 505
DB 25606 CACGTGAGTGAGGCTCTGCGCGCTCTGCGCGGAGAGGCGCGCCCTCTCTGAC 25665
QY 506 TGATCCGAACCCAGAGGCGCGGCTGTCTGATGAGGCGGCGCTCCGCGGTGACCGG 565
DB 25666 CCGCTGACGTGTACACCGCGCTCTTTCGCTGATGATGCTTCTGCGCGCTCTG 25725
QY 566 CCGAG-TGCTGTGATGAGCGCGGAGCGCGGCTCTGACAGCGAGCGCGCATGCGCAC 624
DB 25726 CGCTGCTCGGGGTAGAGCGCGCGCGCTCTGCGGACATGAGGCGAGATGCGCGC 25785
QY 625 GCATGCGCGGAGCGCTTACGCTTACATCAATCAATGACAACTTGGCAACTGAC 684
DB 25786 GCCTTCTGCGAGGCGCTCTCTCTGAGAGCGCGCGCATGCGCGCGCTCTGCGAGC 25845
QY 685 GCGAGTCTGCGCGCGGATCCACACTGCTACTCATGCGGCTGACAGCTGAGGTCG 744
DB 25846 AAGGCTTACACCGCTGCGCGGACAGGCGCATGCGCGCTGAGCTGCGCGCTCT 25905
QY 745 GTCAAACTGCGGACCTGTGATTTGGGCGCTCTCTGTGCGGAGCGCC 792
DB 25906 GACCTCCAGACCTACTGCTCTCTGCGGAGCGAGCGCTCTCATCTGCC 25953

```

RESULT 8

```

US-08-457-335A-6
: Sequence 6, Application US/08457335A
: Patent No. 5723759
: GENERAL INFORMATION:
: APPLICANT: Schupp, Thomas
: APPLICANT: Ligon, James M.
: APPLICANT: Beck, James Joseph
: APPLICANT: Hill, Dwight Steven
: APPLICANT: Ryals, John Andrew
: APPLICANT: Gaffney, Thomas Deane
: APPLICANT: Lam, Stephen Ting
: APPLICANT: Hammer, Phillip B.
: APPLICANT: Ukens, Scott Joseph
: TITLE OF INVENTION: Genes for the synthesis of
: NUMBER OF SEQUENCES: 22
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Ciba-Geigy Corporation
: STREET: 7 Skyline Drive
: CITY: Hawthorne
: STATE: NY
: COUNTRY: USA
: ZIP: 10532
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/457,335A
: FILING DATE: 01-JUN-1995
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/457,205
: FILING DATE: 01-JUN-1995
: APPLICATION NUMBER: 08/258,261
: FILING DATE: 08-Jun-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Elmer, James Scott
: REGISTRATION NUMBER: 36,129
: REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
: TELECOMMUNICATION INFORMATION:

```


Db 25726 GCCTGCTGCGGGTATAGACCCCGCCGCTGCGGCCACAGTCAAGGCGAGATCGCCGC 25785
 Oy 625 GGCAATGGGCGGCACCGCTTACGGTTCTAGACATCAATCAATCAAACTTGGCAACTGAC 684
 Db 25786 GCTTCTGCTGCGAGGCGCTCTCTCTCCCTCGAGAGCGCGCCCGCATCGCCGCTGGCGACG 25845
 Oy 685 GCCAGTCTCTGCGGCGGATCCACATCGCTACTCATCGGCTTACGAGCTGAGAGGTGCG 744
 Db 25846 AAGGCGTCTACACACCGTGGCGCGCAACGGGGCCATGGGCGCGCTGAGACTGGGGCCCTCC 25905
 Oy 745 GTCAAACTGCGGACCTGTGATTGGGCGCTCTGCTGCTGCCAGGCGCC 792
 Db 25906 GACCTTCAGACCTCACTCGCTCCCTGCGGCGACAGGCTCTCATGCGC 25953
 RESULT 10
 US-08-764-233A-1
 : Sequence 1, Application US/08764233A
 Patent No. 5716849
 : GENERAL INFORMATION:
 APPLICANT: Ligon, James M.
 APPLICANT: Schupp, Thomas
 APPLICANT: Beck, James J.
 APPLICANT: Hill, Dwight S.
 APPLICANT: Neff, Snezana
 APPLICANT: Ryals, John A.
 TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Ciba-Geigy Corporation
 STREET: 520 White Plains Road, P.O. Box 2005
 CITY: Tarrytown
 STATE: NY
 COUNTRY: USA
 ZIP: 10591
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/764,233A
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/729,214
 FILING DATE: 09-OCT-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/258,261
 FILING DATE: 08-JUN-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Meigs, J. Timothy
 REGISTRATION NUMBER: 36,241
 REFERENCE/DOCKET NUMBER: 1506/CIP6
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (919) 541-8587
 TELEFAX: (919) 541-8689
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS: 1:
 LENGTH: 49377 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 ORIGINAL SOURCE:
 ORGANISM: Sorangium cellulosum
 IMMEDIATE SOURCE:
 CLONE: p98/1, pUL3, and pYKM15
 FEATURE:
 NAME/KEY: misc.feature
 LOCATION: 383..760
 OTHER INFORMATION: /product= "SorR"
 OTHER INFORMATION: /note= "This gene encodes a protein that is highly homologous

```

1 OTHER INFORMATION: the reductase domains of type I PRSS such as eryA from
2 OTHER INFORMATION: Saccharopolyspora erythraea."
3 FEATURE:
4 NAME/KEY: misc_feature
5 LOCATION: 927..19874
6 OTHER INFORMATION: /product= "Sora"
7 OTHER INFORMATION: /note= "Gene product is highly homologous to type I PRSS t
8 OTHER INFORMATION: are known to be involved in the synthesis of polyketide
9 OTHER INFORMATION: compounds."
10 FEATURE:
11 NAME/KEY: misc_feature
12 LOCATION: 942..7115
13 OTHER INFORMATION: /product= "Module 1 of Sora"
14 FEATURE:
15 NAME/KEY: misc_feature
16 LOCATION: 7203..12884
17 OTHER INFORMATION: /product= "Module 2 of Sora"
18 FEATURE:
19 NAME/KEY: misc_feature
20 LOCATION: 13455..19616
21 OTHER INFORMATION: /product= "Module 3 of Sora"
22 FEATURE:
23 NAME/KEY: misc_feature
24 LOCATION: 19871..46318
25 OTHER INFORMATION: /product= "SorB"
26 OTHER INFORMATION: /note= "Gene product is highly homologous to type I PRS ge
27 FEATURE:
28 NAME/KEY: misc_feature
29 LOCATION: 19870..24556
30 OTHER INFORMATION: /product= "Module 1 of SorB"
31 FEATURE:
32 NAME/KEY: misc_feature
33 LOCATION: 24638..30820
34 OTHER INFORMATION: /product= "Module 2 of SorB"
35 FEATURE:
36 NAME/KEY: misc_feature
37 LOCATION: 30881..35446
38 OTHER INFORMATION: /product= "Module 3 of SorB"
39 FEATURE:
40 NAME/KEY: misc_featur
41 LOCATION: 35528..40114
42 OTHER INFORMATION: /product= "Module 4 of SorB"
43 FEATURE:
44 NAME/KEY: misc_feature
45 LOCATION: 40190..46318
46 OTHER INFORMATION: /product= "Module 5 of SorB"
47 FEATURE:
48 NAME/KEY: misc_feature
49 LOCATION: 46851..47891
50 OTHER INFORMATION: /product= "SorM"
51 OTHER INFORMATION: /note= "The protein encoded by the sorM gene is highly
52 OTHER INFORMATION: homologous to the methyltransferase from Streptomyces
53 OTHER INFORMATION: hygroscopicus that is involved in the synthesis of the
54 OTHER INFORMATION: polyketide rapamycin."
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919

```

[illegible]


```

Db 1310 CCCCAGGCGCTTCTCTACACCGAGGCCGTGCTGCGGCTTCATGTGGCGGCCGAGCA 1369
QY 645 GGTTCAGACATCAACATCGACAAACTTCGCAACTCGACGCCGAGTTCTGCGCCGGAT 704
Db 1370 CGGCGTCGACGTGACCAACAACACTATTACACCGACCCGCTGCTACTTCAACTGCAAGGA 1429
QY 705 CCACACTCGGTACTCATCGGCGTACGAGCTCGAGGTGCGCTCAAAAGCTGCGACTGGT 764
Db 1430 CGACCCCGACCGAGAGGCGCTCTGCGAGGCCGCTCTCGGGGCTCCCGGTACGCGGAGAA 1489
QY 765 GATTGGGCGCGTCCCTGGTGCCAGGCGCC 792
Db 1490 GAAGGGCGCGGTCAACGTGCGCGCGGCC 1517

```

Search completed: June 22, 2000, 15:03:42
 Job time: 17700 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 12:07:14 ; Search time 5541.94 Seconds

(without alignments)
903.245 Million cell updates/sec

Title: US-09-362-485-3

Perfect score: 1235
Sequence: 1 ATCTTCAGATTATATCAGAC.....ANGGAAGCATGATGTCGNC 1235

Scoring table: IDENTITY - NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: gb_gss7:*
93: gb_gss8:*
94: gb_gss9:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: gb_gss10:*
103: gb_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES

Query Match	Best Local Similarity	Matches	Conservative	Mismatches	Indels	Gaps
442	4.6%; Score 57.4; DB 82; Length 925;	43;	176;	142;	0;	0;
920	SCSCSCSBSCSSSSMTSSNSSBSCSSBSBSSSTSSWSSSSBSSSSSSSGTSS	861				
502	CACCTGTGCGAAGCAAGGAGCGCGTGTGCTGATGAGGCGGGGTCGCCGCTCGAA	561				
860	SACVYKCAVSSCCGCCCGGABCCMCSSSSSCGASAGVAVRASGAGRGGSGGASAA	801				
562	CCGCGCAGCTGTGTGTGATCGCCGCCGCAACCGGCGCTCAACGAGCGCCGCTCGCC	621				
800	SHSSSSACBSSSSSCASCMWASSSSSSSASMSRSGGAGGAGSSSSSSSSSSSSASAS	741				
622	AAGCGCATGGGCGCCGACCTTACGCTTTACAGACATCAACATCGAACAATCTGGCAACTC	681				
740	VWSASSSSSSSSSVSCSVASMSCSBSSSSSSASASSSSSSSSASACSCCCTTSMGC	681				
682	GAGCGGAGTTGCGCGCCGACACACTCGCTACATCGGCGCTCGAGCTCGAGGCT	741				

FEATURES	source	Location/Qualifiers
FEATURES	source	Location/Qualifiers
1.846		
/organism="Drosophila melanogaster"		
/plasmid="pBelobAC11"		
/db_xref="taxon:7227"		
/clone_11b="DrosBAC"		
/clone="BACN04N13"		
/note="end : 77"		
BASE COUNT	300 a	47 c 40 g 175 t 284 others
ORIGIN		
Query Match	3.8%; Score 47; DB 82; Length 846;	
Best Local Similarity	14.0%; Pred. 5.7;	
Matches	30; Conservative 11; Mismatches 73; Indels 0; Gaps 0	
QY	995	CGCTGACCAACGAGCAGATGCCGTATGCTCGACGCTTCCGACCATGCTGCGGGCGG 1054
Db	801	SSSSSTSSSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSS 742
QY	1055	CTGCGCGGTGCAATCGGCGACTAGCCAAAGGCTTTCGACGACGACGAGGGCGGTACTGT 1114
Db	741	SASASTSTSTTTGCGBSBSBSTSGTSSSTSBRTBTBTSSBSBSBSSTSTSTBTB 682
QY	1115	CCGACGCGGTGCGCAAGCAGCTGGGGGCGCCGTTCACCGA TCCCGCAGCGTGGTGGCT 1174
Db	681	SSSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSS 622
QY	1175	GACCTCGGGCGCGTGTACGCCGACGACACACATC 1208
Db	621	SSSSCBTSSCSSSSSCCTCGCCSCCCCTC 588
RESULT	7	
CNS0181E/c		
LOCUS		
DEFINITION	CNS0181E 1101 bp DNA GSS 26-JUL-1995	
ACCESSION	Drosophila melanogaster genome survey sequence SP6 end of BAC	
VERSION	BACN3/H05 of DrosBAC library from Drosophila melanogaster (fruit	
KEYWORDS	fly), genomic survey sequence.	
SOURCE	AL108764	
ORGANISM	AL108764.1 GI:5629068	
REFERENCE	GSS.	
AUTHORS	fruit fly.	
TITLE	Drosophila melanogaster	
JOURNAL	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	
COMMENT	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;	
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
	1 (bases 1 to 1101)	
	Genoscope.	
	Direct Submission	
	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :	
	BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr	
	- Web : www.genoscope.cns.fr)	
	Determination of this BAC-end sequence was carried out as part of a	
	collaboration with the European Drosophila Genome Project (EDGP) -	
	http://www.edgp.abl.ac.uk - This Drosophila melanogaster BAC	
	library (Dros BAC) was made by Alain Billand at CSHB (Centre	
	d'Etude du Polymorphisme Humain) with funding provided by a MRC	
	project grant. The DNA was prepared from embryos by Alain Bucheton	
	and Genevieve Payan. It has been constructed in the vector	
	pBelobAC11.	
FEATURES	source	Location/Qualifiers
1.1101		
/organism="Drosophila melanogaster"		
/plasmid="pBelobAC11"		
/db_xref="taxon:7227"		
/clone_11b="DrosBAC"		
/clone="BACN3/H05"		
/note="end : SP6"		
BASE COUNT	61 a 215 c 301 g 235 t 289 others	
ORIGIN		

Query Match	3 %	Score 46.6	DB 83	Length 1101
Best Local Similarity	40.4%	Pred. No. 7.1		
Matches 101	Conservative 40	Missmatches 109	Indels 1	Gaps 1
Db	389	CGGCACACGATTCGCTACGAGACCGCTGACGACCGGACGAGCGGACACACCCCTGC	448	
Db	308	CCCCACVYMYVGMMAASCCCAVASCSCSCSCSCSCSCSCSCSCSCSCSCSCSCG	249	
QY	449	TTGCCCCGATGACGAATGCGCGGCTGCACCTCGCCCGCGAGTTGGCGTTACCACTGA	508	
Db	248	AAGMGSMGSGVGCACAAAGGCGGAGACCGGCCACSCCSCSSGCGSCSSCSCSSSS	189	
QY	509	TGCGAACCCACAGAGGGGCGCGGTGTCTATATGGCGGGGTGCTCCGCGCTCAACCGCCG	568	
Db	188	CGCGACCCCSAGGGGGCCCSGCGGCCACCCCGCGGACVACCGCCAC-CMVGGGGG	130	
QY	569	ACGTGTGTGTATCGGCGCCJGACCCGCGCGGTACACGACCGCCGATCGCCACGGA	628	
Db	129	ACGCAAAAGGAGAGAGCGCGGAGAGCCCGCCCGCCACCCCGCCCGCCACGCGCGGGA	70	
QY	629	TGGCGCGGACG 639		
Db	69	AGCGCCCAAC 69		
RESULT 8				
CNS006XK				
LOCUS	935 bp	DNA	GSS	03-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC #			
LOCUS	BACR14N09 of RPCT-98 library from drosophila melanogaster (fruit fly), genomic survey sequence.			
ACCESSION	AL066051.1	GI:4945019		
VERSION	GSS.			
KEYWORDS	fruit fly.			
SOURCE	Drosophila melanogaster			
ORGANISM	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
REFERENCE	1 (bases 1 to 935)			
AUTHORS	Genoscope.			
TITLE	Direct Submission			
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr			
COMMENT	- Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCT-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .			
FEATURES	Location/Qualifiers			
SOURCE	1..935			
	/organism="Drosophila melanogaster"			
	/db_xref="taxon:7227"			
	/clone_lib="RPCT-98"			
	/clone="BACR14N09"			
	/note="end : 77"			
BASE COUNT	257 a 170 c 162 g 96 t 250 others			
ORIGIN				

Best Local Similarity 27.7%; Pred. No. 7.5; Matches 77; Conservative 76; Mismatches 125; Indels 0; Gaps 0;

368 GCACGATGCGTGTGATTCGGGACGACGATTCCTAGAGAGACGCTCAGACCG 427
 511 SCMCRTSSSTGCTCCTCMTSSSVSCGCTCCTCCTCKMCTCTYCKCGCGC 570
 428 CGCAGCGGACATACCCCTCTGCCCCGATGAGCGAAGTGGCGGCTCGACTCGCC 487
 571 CSTSSSSSCBSYSTCCTCTCKTSSGCTGCTCCGGGAGCGGCGCGCGC 630
 488 AGTGTGCGCTTACCACTGATCGAAGCGAAGGCGCGCGCTGCTGATGCGGGG 547
 631 SGGGCGGCGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSG 690
 548 TGCCCGCGCTGCAACCGCGCGCGCTGATGCGCGCGCGCGCGCGCTACACG 607
 691 GCGCGSSGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 750
 608 CAGCCCGCATCGCCACGCGATGGCGCGCGCTACG 645
 751 SGGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 788

RESULT 9
 CNS01213 645 bp DNA GSS 26-JUL-1999
 LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC
 DEFINITION BACN08C07 of Drosophila library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 ACCESSION AL101589
 VERSION AL101589.1 GI:5613200
 KEYWORDS GSS.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 645)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage;
 BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT Determination of this BAC-end sequence was carried out as part of a
 collaboration with the European Drosophila Genome Project (BDGP) -
 http://www.edgp.edi.ac.uk - This Drosophila melanogaster BAC
 library (Dros BAC) was made by Alain Billard at CEPH (Centre
 d'Etude du Polymorphisme Humain) with funding provided by a MRC
 project grant. The DNA was prepared from embryos by Alain Bucheton
 and Genevieve Payan. It has been constructed in the vector
 pBelobac11.

FEATURES
 Location/Qualifiers
 1..645
 /organism="Drosophila melanogaster"
 /plasmid="pBelobac11"
 /db_xref="taxon:7227"
 /clone_lib="DrosBAC"
 /clone="BACN08C07"
 /note="end : 77"

BASE COUNT 28 a 26 c 85 g 92 t 414 others
 ORIGIN

Query Match 3.7%; Score 46.2; DB 82; Length 645;
 Best Local Similarity 10.3%; Pred. No. 7.8; Matches 43; Conservative 135; Mismatches 238; Indels 3; Gaps 1;

311 GCGTGGACAGGCGGACGATCTTTCAGCTTTCGATTTGGCGCGCTACGCTTGA 370
 633 SSSNN 574
 371 CGATCGTGTGATTCGCGACACGATGATCTCTACGAGACGCGTCAACCGCG 430

573 CAGSNNSGNS 514
 431 AGGCGGCTACCCCTGCTTCCCGATGAGCAAGTCCCGCTGACTCGCCCGCAG 490
 513 ATNNNNCTMAANASANTSAANGSNNSNSNSNSNSNSNSNSNSNSNSNS 454
 491 TTGGCGCTTACACCTGATGATCAACCGAGGGGCGCGGCTGTGATGACCGGGTGC 550
 453 NNNNSSTSSAAAASSSSCASNSNSNSNSNSNSNSNSNSNSNSNSNSNSNS 394
 551 CCGCGCTCGAACCGCGCGCGCTGCTGATCG--CGCCGCAACCGCGCGCTCAACG 607
 393 SSSSSSSSCSSCNMSAATSNNSNSNSNSNSNSNSNSNSNSNSNSNSNSNS 334
 608 CAGCCCGCATCGCCACGCGATGGCGCGCGCGCTTACGCTTACATCAATCAGCA 667
 333 SSNSSSCATSNNSASNSANASSSSSNSNSNSNSNSNSNSNSNSNSNS 274
 668 AACTTCGCACTGACGCGGATCTGCGCGCGATCCACACTGCTACTCGGCC 726
 273 NSNSSSSCCNNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNS 215

RESULT 10
 CNS006K/C 935 bp DNA GSS 03-JUN-1999
 LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC #
 DEFINITION BACR14N09 of RPCT-98 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 ACCESSION AL066051
 VERSION AL066051.1 GI:4945019
 KEYWORDS GSS.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 935)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage;
 BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT Determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see http://www.fruitfly.org The BDGP Drosophila
 melanogaster BAC library was prepared by Kazuhiro Oseegawa and
 Aaron Mamoser in Pieter de Jong's laboratory in the Department of
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
 NY. The library is named RPCT-98 and was constructed by partial
 EcoRI digestion of Drosophila DNA provided by the BDGP from the
 isogenic strain y2; cn bw sp, the same strain used for the BDGP's
 P1 and EST libraries. A more detailed description of the library
 and how to order individual BAC clones, the entire library or
 filters for hybridization from the BACPAC Resource Center can be
 found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
 Location/Qualifiers
 1..935
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone_lib="RPCT-98"
 /clone="BACR14N09"
 /note="end : 77"

BASE COUNT 257 a 170 c 162 g 96 t 250 others
 ORIGIN

Query Match 3.7%; Score 45.8; DB 82; Length 935;
 Best Local Similarity 27.7%; Pred. No. 9.8; Matches 94; Conservative 83; Mismatches 162; Indels 0; Gaps 0;


```

FEATURES
  source      Location/Qualifiers
              1..744
              /organism="Drosophila melanogaster"
              /plasmid="pBelobAC11"
              /db_xref="taxon:7227"
              /clone_lib="DrosBAC"
              /clone="BACN37B10"
              /note="end : Sp6"
BASE COUNT   20 a 191 c 245 g 105 t 183 others
ORIGIN
Query Match      3.7%; Score 45.2; DB 83; Length 744;
Best Local Similarity 34.0%; Pred. No. 12;
Matches 87; Conservative 41; Mismatches 128; Indels 0; Gaps 0;

QY 329 TCTGTTCAGCTTCTTCAATTTGGCCGCTCAGCTGCTTGCACGATGCTTGTGGATT 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 191 YCTTGTGCGCTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 250

QY 389 CCGGCACACACTCAATTCCTACGACGCTCCAGACCGCGCGCGCGCGCGCGCGCGCG 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 251 YCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 310

QY 449 TTGCCCCGATGACGAGTCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 311 CGGCCCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 370

QY 509 TCGGACCCAGAGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 371 CCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 430

QY 569 ACGTCGTGTCATCGG 584
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 431 NGGNGSGSGCTKTGCG 446

RESULT 13
LOCUS      CNS0175V 1101 bp DNA GSS 26-JUL-1999
DEFINITION BACN37108 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION  AL108460
VERSION     AL108460.1 GI:5628764
KEYWORDS   GSS.
SOURCE     fruit fly.
           Drosophila melanogaster
           Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
           Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
           Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
           1 (bases 1 to 1101)
REFERENCE   Direct Submission
           Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
           BP 191 91006 EVRY cedex - FRANCE (E-mail : sequefgenoscope.cns.fr
           - Web : www.genoscope.cns.fr)
           Determination of this BAC-end sequence was carried out as part of a
           collaboration with the European Drosophila Genome Project (EDGP) -
           http://www.edgp.ebi.ac.uk/. This Drosophila melanogaster BAC
           library (Dros BAC) was made by Alain Billaut at CEPH (Centre
           d'Etude du Polymorphisme Humain) with funding provided by a MRC
           project grant. The DNA was prepared from embryos by Alain Bucheton
           and Genevieve Payan. It has been constructed in the vector
           pBelobAC11.
FEATURES
  source      Location/Qualifiers
              1..1101
              /organism="Drosophila melanogaster"
              /plasmid="pBelobAC11"
              /db_xref="taxon:7227"
              /clone_lib="DrosBAC"
              /clone="BACN37108"
              /note="end : Sp6"

```

```

BASE COUNT   254 a 176 c 160 g 152 t 359 others
ORIGIN
Query Match      3.7%; Score 45.2; DB 83; Length 1101;
Best Local Similarity 12.6%; Pred. No. 13;
Matches 40; Conservative 157; Mismatches 121; Indels 0; Gaps 0;

QY 420 CCAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 479
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 755 CGASAGGSGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 814

QY 480 CGCCCGCGAGTTGGCGCTTACACCTGATCGAACCCAGGGGCGCGCGCGCGCGCG 539
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 815 CSAVASASASVMSKVASAVASCAVASAGAVSSCRSVASVSAASVSSVSS 874

QY 540 GGGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 599
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 875 SSSVVSAAAVASASSSSASMAVAAAAVAVSSVSVASVSSSSSSSSSVASVSA 934

QY 600 CTACACCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 659
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 935 SVAAASASSVSSSSSVSTSSASVSAVSAVSSASVSSSVSVSVAAVAA 994

QY 660 CATGACAACTTGGCGCACTGACCGCGAGTTGCGGCGCGGATCCACACTGCTACT 719
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 995 AAAAAAASASSSSASAAVAVASASSSSSSSSSSSSSSSSSSSSSVSVSVSS 1054

QY 720 ATCGGCTACGAGCTCGA 737
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1055 VSVAAVSAASASVSVSA 1072

```

```

RESULT 14
LOCUS      A1712257 315 bp mRNA EST 02-FEB-2000
DEFINITION 605069D09.x1 605 - Endosperm cDNA library from Schmidt lab Zea mays
ACCESSION  A1712257
VERSION     A1712257.1 GI:5006195
KEYWORDS   EST.
SOURCE     Zea mays.
           Zea mays
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
           Poaceae; Zea.
           1 (bases 1 to 315)
REFERENCE   1 (bases 1 to 315)
           Walbot V.
           Maize ESTs from various cDNA libraries sequenced at Stanford
           University
           Unpublished (1999)
           On Jun 22, 1998 this sequence version replaced gi:3246668.
           Contact: Walbot V
           Department of Biological Sciences
           Stanford University
           855 California Ave, Palo Alto, CA 94304, USA
           Tel: 650 723 2227
           Fax: 650 725 8221
           Email: walbot@stanford.edu
           Plate: 605069 row: D column: 09.
FEATURES
  source      Location/Qualifiers
              1..315
              /organism="Zea mays"
              /cultivar="Ohio43"
              /db_xref="taxon:4577"
              /clone_lib="605 - Endosperm cDNA library from Schmidt lab"
              /tissue_type="nucellar, embryo, and endosperm"
              /dev_stage="10-14 days post-pollination"
              /lab_host="DH5(alpha)"
              /note="Organ: kernel; Vector: pMD-GAL4-2'; Site:1: EcoRI;
              Site:2: XhoI; kernel endosperm cDNA library from Schmidt
              lab"
BASE COUNT   61 a 105 c 106 g 42 t 1 others

```


ORIGIN

Query Match 3.6%: Score 45; DB 51; Length 315;

Best Local Similarity 50.7%: Pred. No. 12; Matches 108; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 391 GGCACACGTCATTCCTACGAGACCGTCACAGCCGCGCAGTACCCCTGCTT 450
 DB 58 GGCATGGCGCTCCATTCACAAAGACAGAGTGGTGGCCAGCAACTTAAACACACCCC 117
 QY 451 GCGCCCATGAGCAATGCGCGGCTGCAGTCCGCCAGGTTGGCCTTACACCTGATG 510
 DB 118 GCGCTTAAGCAGCCAGCGCGGTCAAGGATAGCGCGCGCGGTCCGAGCCAGC 177
 QY 511 CGAACCCAGAGGCGCGCGGTGCTGATGGCGGGGTGCCCGCTCAACCGCGCGAC 570
 DB 178 CGTGGGACAGGGGCGCGCGGTCTTCTGTAGTCTTCTGTCGTGCGCGCGGAGGC 237
 QY 571 GTGCTGTATCGCGCGCGCGCGCGCTAC 603
 DB 238 GTGGCGGTGGCGCGCGCGCGCTCC 270

RESULT 15

CNS010RJ

LOCUS 846 bp DNA GSS 26-JUL-1999
 DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC

ACCESSION BACN04N13 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
 VERSION AL099337
 KEYWORDS GSS: AL099337.1 GI:5610948

SOURCE

ORGANISM fruit fly.
 Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 846)

REFERENCE

AUTHORS Genoscope.

TITLE Direct Submission
 JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)

COMMENT

Determination of this BAC-end sequence was carried out as part of a
 collaboration with the European Drosophila Genome Project (EDGP) -
 http://www.edgp.edl.ac.uk -. This Drosophila melanogaster BAC
 library (Dros BAC) was made by Alain Billand at CEPH (Centre
 d'Etude du Polymorphisme Humain) with funding provided by a MRC
 project grant. The DNA was prepared from embryos by Alain Bucheton
 and Genevieve Payan. It has been constructed in the vector
 pBelOBAC11.

FEATURES

Location/Qualifiers
 1..846
 /organism="Drosophila melanogaster"
 /plasmid="pBelOBAC11"
 /db_xref="taxon:7227"
 /clone_lib="DrosBAC"
 /clone="BACN04N13"
 /note="end : T7"

BASE COUNT 300 a 47 c 40 g 175 t 284 others
 ORIGIN

Query Match 3.6%: Score 45; DB 82; Length 846;

Best Local Similarity 16.2%: Pred. No. 14; Matches 38; Conservative 109; Mismatches 88; Indels 0; Gaps 0;

QY 506 TGATGGCAACCAAGGCGCGGTGTGTGATGGCGGCGCGCGGTGGAACCG 565
 DB 574 KGAGCGRAAASSAAGAGGGGCGCGAGCGSSSSSSSSSSSSSSSSSSSS 633
 QY 566 CCGAGCTGCTGTGATGGCGCGCGCGCGCTACAAAGCAGCCCGCATGCCAAG 625

DB 634 SSGSSSSAAASAAASVSSASASVSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 693
 QY 626 GCATGGCGCGCGCGCGGTTCAGGTTCTAGACATCAACATGACAACTTGGCACTCGACG 685
 DB 694 SSTSVSSSSAAVAVAVSSASASVSSSVGCAASSASASTSTSSSSSSSSSSG 753
 QY 686 CCGAGTTCTGCGCGCGCGATCCACACTCGCTACTCATCGGCTACGAGCTCGAGG 740
 DB 754 SSASGSSSSAAASAAASVSSASASVSSSSSSSSSSSSSSSSSSSSSSSSSSSS 808

Search completed: June 22, 2000, 12:07:22
 Job time: 10356 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 14:51:53 ; Search time 8627.09 Seconds

(Without alignments)
-139,484 Million cell updates/sec

Title: US-09-362-485-4

Perfect score: 1237
Sequence: 1 ATCTTGACAGTTATTCGAC.....GGAGCGGATGTCGCGC 1237

Scoring table: IDENTITY_MDC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

Genembi:*

1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pi1:*
8: gb_pi2:*
9: gb_pi3:*
10: gb_pi4:*
11: gb_pi5:*
12: gb_pi6:*
13: gb_pi7:*
14: gb_pi8:*
15: gb_pi9:*
16: gb_pi10:*
17: gb_pi11:*
18: gb_pi12:*
19: gb_pi13:*
20: gb_pi14:*
21: gb_pi15:*
22: gb_pi16:*
23: gb_pi17:*
24: gb_pi18:*
25: gb_pi19:*
26: gb_pi20:*
27: gb_pi21:*
28: gb_pi22:*
29: gb_pi23:*
30: gb_pi24:*
31: gb_pi25:*
32: gb_pi26:*
33: gb_pi27:*
34: gb_pi28:*
35: gb_pi29:*
36: gb_pi30:*
37: gb_pi31:*
38: gb_pi32:*
39: gb_pi33:*
40: gb_pi34:*
41: gb_pi35:*
42: gb_pi36:*
43: gb_pi37:*
44: gb_pi38:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1237	100.0	1237	5 A87607	A87607 Sequence 4
2	1237	100.0	1237	5 A89747	A89747 Sequence 4
3	1237	100.0	56414	1 MTVO02	AL008967 Mycobacte
4	1234.4	99.8	1236	5 A87613	A87613 Sequence 10
5	1234.4	99.8	1236	5 A89753	A89753 Sequence 10
6	1230.4	99.5	1235	5 A87605	A87605 Sequence 3
7	1230.4	99.5	1235	5 A89746	A89746 Sequence 3
8	1228	99.3	1228	5 A87608	A87608 Sequence 5
9	1228	99.3	1228	5 A89748	A89748 Sequence 5
10	1221.4	98.7	1235	5 A89749	A89749 Sequence 6
11	1221.4	98.7	1235	5 A87611	A87611 Sequence 8
12	1220.8	98.7	1235	5 A89751	A89751 Sequence 8
13	1220.8	98.7	1235	5 A87605	A87605 Sequence 2
14	1219.4	98.6	1245	5 A89745	A89745 Sequence 2
15	1219.4	98.6	1260	5 A87604	A87604 Sequence 1
16	1219.4	98.6	1260	5 A89744	A89744 Sequence 1
17	1219.4	98.6	1260	5 A89744	A89744 Sequence 1
18	1219.4	98.6	1260	5 A89744	A89744 Sequence 1
19	1213.8	98.1	1229	5 A87610	A87610 Sequence 7
20	1213.8	98.1	1229	5 A89750	A89750 Sequence 7
21	1206.4	97.5	1209	5 A87612	A87612 Sequence 9
22	1199.4	97.0	1208	5 A89752	A89752 Sequence 9
23	1192.4	96.4	1194	2 MTU92472	U92472 Mycobacteri
24	698	56.4	720	5 A89759	A89759 Sequence 16
25	538.4	43.5	40745	1 SC151	AL109848 Streptomy
26	467.6	37.8	682	5 A89758	A89758 Sequence 15
27	467.6	37.8	682	5 A89758	A89758 Sequence 15
28	377	30.5	8046	2 AF049107	AF049107 Myxococc
29	351.2	28.4	11514	2 AE002028	AE002028 Deinococc
30	331.4	26.8	2780	1 RLEZ38118	AJ238118 Rhizobium
31	330.8	26.7	1585	2 AF070716	AF070716 Vibrio pr
32	312.8	25.3	1736	1 BACALDHC	M33399 B. stearothe
33	293	23.7	1477	2 AF070715	AF070715 Shewanell
34	288.4	23.3	2113	1 D37807	D37807 Phormidium
35	277.2	22.4	1125	5 E01847	E01847 DNA encodin
36	258	20.9	137740	1 D90900	D90900 Synecocyst
37	243.2	19.7	1776	1 AB013821	AB013821 Enterobac
38	243	19.6	1776	1 BACALDH	M33398 B. sphaericu
39	235.6	19.0	1709	1 BACALD	L20916 Bacillus su
40	235.6	19.0	12316	1 B5820215	B5820215 Bacillus
41	235.6	19.0	217420	1 B5820215	B5820215 Bacillus
42	218.2	17.6	1714	2 AF070714	AF070714 Carnobact
43	107.2	8.7	10552	2 AE001565	AE001565 Helicobac
44	99.2	8.0	10512	2 AE000639	AE000639 Helicobac
45	74.2	6.0	3160	1 RUC05294	U05294 Rhodospirill

ALIGNMENTS

RESULT 1
 LOCUS A87607 1237 bp DNA
 DEFINITION Sequence 4 from Patent WO9836089.
 ACCESSION A87607
 VERSION A87607.1 GI:6736247
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1237)
 AUTHORS Flohe, L. and Singh, M.
 TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
 JOURNAL Patent: WO 9836089-A 20-AUG-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source 1..1237
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 236 a 394 c 386 g 221 t
 ORIGIN

Query Match 100.0%; Score 1237; DB 5; Length 1237;
 Best Local Similarity 100.0%; Pred. No. 2, 3e-170;
 Matches 1237; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTTGACGATTATTCGAACTTTCTTCACTGAAAGCTACAGTATCGAGGGGTAATC 60
 DB 1 ATCTTGACGATTATTCGAACTTTCTTCACTGAAAGCTACAGTATCGAGGGGTAATC 60
 QY 61 ATGCGCGTGGTATTCGACGAGACGACAAAACAAAGAAATTCGGGAGGATACCCCG 120
 DB 61 ATGCGCGTGGTATTCGACGAGACGACAAAACAAAGAAATTCGGGAGGATACCCCG 120
 QY 121 GCCGCGTGGGAACTAACCCGTGCGGATGAGTGTCTATCCAGCAGGTGCCGA 180
 DB 121 GCCGCGTGGGAACTAACCCGTGCGGATGAGTGTCTATCCAGCAGGTGCCGA 180
 QY 181 GAGGCGTGGGATATCACGAGCGGATTCAGAGCGGCGGCAACTGTGGGCAAC 240
 DB 181 GAGGCGTGGGATATCACGAGCGGATTCAGAGCGGCGGCAACTGTGGGCAAC 240
 QY 241 GCCGACGAGTGTGGGCGGACGCTGATTTATGCTCAAGTCAAGAAACGATAGCGG 300
 DB 241 GCCGACGAGTGTGGGCGGACGCTGATTTATGCTCAAGTCAAGAAACGATAGCGG 300
 QY 301 GAATACGCGCGGCTGCGACAGGCGCAATCTTTTCACGTTCTTGCAATTGGCGGCTCA 360
 DB 301 GAATACGCGCGGCTGCGACAGGCGCAATCTTTTCACGTTCTTGCAATTGGCGGCTCA 360
 QY 361 CGTCTTGACACGATGCGTGTGATTCGCGACACGATTCGCTTACGAGACCGTGC 420
 DB 361 CGTCTTGACACGATGCGTGTGATTCGCGACACGATTCGCTTACGAGACCGTGC 420
 QY 421 CAGACCGCGGACGCGCACTACCCCTGCTTGCCTCCCGATGACGCAAGTGGCGTGAATC 480
 DB 421 CAGACCGCGGACGCGCACTACCCCTGCTTGCCTCCCGATGACGCAAGTGGCGTGAATC 480
 QY 481 GCGCGCGAGGTGGGCTTACCACTGATGAGAACCCCAAGGCGGCGCGGTGTCTATG 540
 DB 481 GCGCGCGAGGTGGGCTTACCACTGATGAGAACCCCAAGGCGGCGCGGTGTCTATG 540
 QY 541 GCGCGGAGTGGCGGCTGTCGACCGCGGACGCTGCTGATTCGCGCGCGGACCGCGG 600
 DB 541 GCGCGGAGTGGCGGCTGTCGACCGCGGACGCTGCTGATTCGCGCGCGGACCGCGG 600
 QY 601 TACAAACGACCGCGCATTCGCAACGCGATGGCGCGACCGTTACGTTACGATCAAC 660
 DB 601 TACAAACGACCGCGCATTCGCAACGCGATGGCGCGACCGTTACGTTACGATCAAC 660
 QY 661 ATGACAAACTTGGCAACTGACGCGAGTTCGCGGCGGATCCACACTCCCTACTCA 720

DB 661 ATGACAAACTTGGCAACTGACGCGAGTTCGCGGCGGATCCACACTCCCTACTCA 720
 QY 721 TCGGCTTACGAGCTCGAGGCTGCCGTCAAAAGTGGCGACCTGATTTGGGCGCTCTG 780
 DB 721 TCGGCTTACGAGCTCGAGGCTGCCGTCAAAAGTGGCGACCTGATTTGGGCGCTCTG 780
 QY 781 GTGCGAGCGCGCAAGGACCAATTAATGCTCGAATTCCTTCCCGCATATGAACA 840
 DB 781 GTGCGAGCGCGCAAGGACCAATTAATGCTCGAATTCCTTCCCGCATATGAACA 840
 QY 841 GTGCGGCTAGTGGTATATAGCATCGACAGGCGGCTGTTTCGAAGGCTCAGACCG 900
 DB 841 GTGCGGCTAGTGGTATATAGCATCGACAGGCGGCTGTTTCGAAGGCTCAGACCG 900
 QY 901 ACACCTACGACCAACCCGAGCTTCGCGCTGACGACACGCTGTTTACGTGGCGAAG 960
 DB 901 ACACCTACGACCAACCCGAGCTTCGCGCTGACGACACGCTGTTTACGTGGCGAAG 960
 QY 961 ATGCGCGCTCGGTGCGGAAGAGCTGACCTACGCGCTGACCAACGCGAGATGCCAT 1020
 DB 961 ATGCGCGCTCGGTGCGGAAGAGCTGACCTACGCGCTGACCAACGCGAGATGCCAT 1020
 QY 1021 GTGCTGAGCTTGGCGACCATATGCTGCGGCGGCGCTGCCGTGCAATCCGCACTAGCC 1080
 DB 1021 GTGCTGAGCTTGGCGACCATATGCTGCGGCGGCGCTGCCGTGCAATCCGCACTAGCC 1080
 QY 1081 AAGGCTTTTCAGCGACGACGAAGGCGGCTTACTGTCCGAAGGGGTGGCCACGACCTGGGG 1140
 DB 1081 AAGGCTTTTCAGCGACGACGAAGGCGGCTTACTGTCCGAAGGGGTGGCCACGACCTGGGG 1140
 QY 1141 GTGCGCTTACCGACGCGCGGCGACGCTGCTGCGCTGACTCTGGCGCGCTGTTAGCGCGAG 1200
 DB 1141 GTGCGCTTACCGACGCGCGGCGACGCTGCTGCGCTGACTCTGGCGCGCTGTTAGCGCGAG 1200
 QY 1201 CACACGTCGAGGATTAAGGAAGCATGATGTGGCGCG 1237
 DB 1201 CACACGTCGAGGATTAAGGAAGCATGATGTGGCGCG 1237

RESULT 2
 LOCUS A89747 1237 bp DNA
 DEFINITION Sequence 4 from Patent WO9832862.
 ACCESSION A89747
 VERSION A89747.1 GI:6738281
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1237)
 AUTHORS Flohe, L. and Singh, M.
 TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
 JOURNAL Patent: WO 9832862-A 30-JUL-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source 1..1237
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 236 a 394 c 386 g 221 t
 ORIGIN

Query Match 100.0%; Score 1237; DB 5; Length 1237;
 Best Local Similarity 100.0%; Pred. No. 2, 3e-170;
 Matches 1237; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTTGACGATTATTCGAACTTTCTTCACTGAAAGCTACAGTATCGAGGGGTAATC 60
 DB 1 ATCTTGACGATTATTCGAACTTTCTTCACTGAAAGCTACAGTATCGAGGGGTAATC 60
 QY 61 ATGCGCGTGGTATTCGACGAGACGACAAAACAAAGAAATTCGGGAGGATACCCCG 120

```

Db 61 ATGGCGCTGCGATTCCGACGAGACCAAAAACAAGATTCCGGGTGGCCATCACCCCG 120
Qy 121 GCCGGCGTGGCGGAATTAACCCGTCGTGGCCATGAGTGTCTATCTACAGGAGTGGCGGA 180
Db 121 GCCGGCGTGGCGGAATTAACCCGTCGTGGCCATGAGTGTCTATCTACAGGAGTGGCGGA 180
Qy 181 GAGGCGTGGCTATTCACGAGCGGATTTCAAGGCGGACGGCGCAACTGTGTGGCACC 240
Db 181 GAGGCGTGGCTATTCACGAGCGGATTTCAAGGCGGACGGCGCAACTGTGTGGCACC 240
Qy 241 GCCGACGAGTGTGGCGGAGCGTATTTATTTGCTCAAGGTCAAGAACCGATAGCGGCG 300
Db 241 GCCGACGAGTGTGGCGGAGCGTATTTATTTGCTCAAGGTCAAGAACCGATAGCGGCG 300
Qy 301 GAATACGCGCGCTGCGACACGCGGAGATCTTTGTCACGTTCTTGCAATTTGGCCGCTCA 360
Db 301 GAATACGCGCGCTGCGACACGCGGAGATCTTTGTCACGTTCTTGCAATTTGGCCGCTCA 360
Qy 361 CGTGGTTCACGAGTATTCGCTTTGATTCGCGACACGCTCAATTTGCTACGAGACGCTC 420
Db 361 CGTGGTTCACGAGTATTCGCTTTGATTCGCGACACGCTCAATTTGCTACGAGACGCTC 420
Qy 421 CAGACCGCGGAGCGGAGCTACCTCCCTGCTGCGGATGAGCGAAGTGGCGGCTGACTC 480
Db 421 CAGACCGCGGAGCGGAGCTACCTCCCTGCTGCGGATGAGCGAAGTGGCGGCTGACTC 480
Qy 481 GCCGCGCGAGTGTGGCGCTTACCACTGATGGAAACCCAAAGCGCGGCTGTCTGATG 540
Db 481 GCCGCGCGAGTGTGGCGCTTACCACTGATGGAAACCCAAAGCGCGGCTGTCTGATG 540
Qy 541 GCGCGGCGCGCGCGCTGGAACCGGCGGAGCTGATGGAGTGGCGGCGGCGGCGGCGGCG 600
Db 541 GCGCGGCGCGCGCGCTGGAACCGGCGGAGCTGATGGAGTGGCGGCGGCGGCGGCGGCG 600
Qy 601 TACAACGAGCGCGCGATGCGCAAGCGGATGGCGGCGGAGCTTACGCTTACGATCAAC 660
Db 601 TACAACGAGCGCGCGATGCGCAAGCGGATGGCGGCGGAGCTTACGCTTACGATCAAC 660
Qy 661 ATGCAACAACTTGGCGCACTGAGCGCGGATTTGCGCGGAGTCCACACCTGCTACTCA 720
Db 661 ATGCAACAACTTGGCGCACTGAGCGCGGATTTGCGCGGAGTCCACACCTGCTACTCA 720
Qy 721 TCGGCGTACGAGCTGAGGCGTCCGCTCAAAAGTGGCGGAGCTGATTTGGGCGGCTGTG 780
Db 721 TCGGCGTACGAGCTGAGGCGTCCGCTCAAAAGTGGCGGAGCTGATTTGGGCGGCTGTG 780
Qy 781 GTGCGAGCGCGCAAGGCGACCCAAATTAATCTGCAATCTTGTGCGCATATGAAACCA 840
Db 781 GTGCGAGCGCGCAAGGCGACCCAAATTAATCTGCAATCTTGTGCGCATATGAAACCA 840
Qy 841 GGTGGGCTACTGTGGATATAGCATGACGAGGCGGCGCTTCTGAAAGGCTCAGACCG 900
Db 841 GGTGGGCTACTGTGGATATAGCATGACGAGGCGGCGCTTCTGAAAGGCTCAGACCG 900
Qy 901 ACCACTACGACCAACCGGAGCTTGGCGGTGACGACGACGCTGTTTACTGCGTGGCGAAC 960
Db 901 ACCACTACGACCAACCGGAGCTTGGCGGTGACGACGACGCTGTTTACTGCGTGGCGAAC 960
Qy 961 ATGCGCGGCTGGGTCGGAAGAGCTGACGACGCTGACGACCAAGGAGGAGCATCCGAT 1020
Db 961 ATGCGCGGCTGGGTCGGAAGAGCTGACGACGCTGACGACCAAGGAGGAGCATCCGAT 1020
Qy 1021 GTGCTCGAGCTTGGCGACATGCGTGGCGGCGGCGTGGCGGATCCGCGACTAGCC 1080
Db 1021 GTGCTCGAGCTTGGCGACATGCGTGGCGGCGGCGTGGCGGATCCGCGACTAGCC 1080
Qy 1081 AAAGTCTTTGAGAGGCAAGGAGGCGTACTATGTCGAAAGGCGGCGACGACCTGGGG 1140
Db 1081 AAAGTCTTTGAGAGGCAAGGAGGCGTACTATGTCGAAAGGCGGCGACGACCTGGGG 1140
Qy 1141 GTGCGCTTACGAGCGGCGGAGGCTGCTGAGCTGAGCTGCGCGGCTGTTAGCGCGAG 1200
Db 1141 GTGCGCTTACGAGCGGCGGAGGCTGCTGAGCTGAGCTGCGCGGCTGTTAGCGCGAG 1200

```

```

Qy 1201 CACACGTGGGAGTAAGGAAGCATGATGCGGCG 1237
Db 1201 CACACGTGGGAGTAAGGAAGCATGATGCGGCG 1237

```

```

RESULT 3
LOCUS MT002 56414 bp DNA BCT 17-JUN-1998
DEFINITION Mycobacterium tuberculosis H37Rv complete genome; segment 122/162.
ACCESSION AL008967 AL123456
VERSION AL008967.1 GI:3251491
KEYWORDS
SOURCE
ORGANISM
Mycobacterium tuberculosis.
Mycobacterium tuberculosis
Bacteria: Firmicutes: Actinobacteria: Actinobacteridae:
Actinomycetales: Corynebacteriaceae: Mycobacteriaceae;
Mycobacterium.
1 (bases 1 to 56414)

```

```

REFERENCE
AUTHORS Cole,S.T., Brosch,R., Parkhill,J., Garnier,T., Churcher,C.,
Harris,D., Gordon,S.V., Eigmeier,K., Gas,S., Barry III,C.E.,
Tekala,F., Badcock,K., Basham,D., Brown,D., Chillingworth,T.,
Comor,R., Davies,R., Devlin,K., Feltwell,T., Gentles,S.,
Hamlin,N., Holroyd,S., Hornsby,T., Jagels,K., Krogh,A., McLean,J.,
Moule,S., Murphy,L., Oliver,S., Osborne,J., Quail,M.A.,
Najandram,M.A., Rogers,J., Rutter,S., Seeger,K., Skelton,S.,
Squares,S., Squires,R., Sulston,J.E., Taylor,K., Whitehead,S. and
Barrell,B.G.

```

```

TITLE Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence

```

```

JOURNAL Nature 393 (6685), 537-544 (1998)
MEDLINE 98295987
REMARK Erratum: [published erratum appears in Nature 1998 Nov
12;396(6707):190]]

```

```

REFERENCE
AUTHORS Parkhill,J.
TITLE Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium
tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome
Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique
Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,
75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk

```

COMMENT

Notes: Details of M. tuberculosis sequencing at the Sanger Centre are available on the World Wide Web.

(URL: http://www.sanger.ac.uk/Projects/M_tuberculosis/) CDS have been renumbered from the original comsid submissions but the old gene designations are in brackets after the new gene numbers. Gene prediction was based on a Hidden Markov Model of TB genes implemented in TPRase (Krogh) supplemented with visual inspection of positional base preference in codons, especially where there is an increase in the observed/expected third position G + C. CAUTION: In some cases we may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, or ttg) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

FEATURES

source

Location/Qualifiers

misc_feature

source

source

/note="Fragment designated v002. Does not represent a physical clone"

<1..233

source

/organism="Mycobacterium tuberculosis"

/strain="H37Rv"

source

/db_xref="taxon:1773"

gene

/clone="1154"

complement(3..527)

CDS
 /gene="recx"
 complement(3. .527)
 /gene="recx"
 /note="Rv2736c, (MTV002.01c), len: 174 aa. recx, similar to eg. RECX_PSEAE P37860 regulatory protein recx from Pseudomonas aeruginosa (153 aa), fasta scores: opt: 161 z-score: 257.2 E(): 3.0e-07, 30.7% identity in 137 aa overlap. Overlaps and extends CDS from overlapping cosmid MTCY154.16c"
 /codon_start=1
 /transl_table=11
 /product="recx"
 /protein_id="CA15532.1"
 /db_xref="GI:2624257"
 /db_xref="SPTREMBL:O33280"
 /translation="MTVSCPSTSEREQARALCRLTRARSTRALAGOLAKRGY PEDIGNVLDRLAAGLVDDTDFAEQWQSRANAKSKRALAEHLAKVDDVITTT VLGIDGAEGRRAEKLVRARLRREVLIIDGTGEARVSRRLVAMLRGCGTLACEV VIAIAERERRRV"
 complement(493. .2865)
 /gene="recA"
 complement(493. .2865)
 /note="Rv2737c, (MTV002.02c), len: 790 aa. recA, identical to RECA_MYCTU P26345 recA protein (790 aa). Contains self-splicing protein element (intron) from 2294 to 969 (c) similar to Intron II from TR:E332317 (EMBL:Y13030) DNA-directed DNA polymerase (BC 2.7.7.7) from Thermococcus sp. (1829 aa), fasta scores: opt: 81 z-score: 233.2 E(): 6e-06, 24.6% identity in 183 aa overlap. Contains PS00017 ATP/GTP-binding site motif A (P-loop), PS00321 recA signature, and PS00811 protein splicing signature. See Davis et al. (1992) Cell 71(2):201-210"
 /codon_start=1
 /transl_table=11
 /product="recA"
 /protein_id="CA15533.1"
 /db_xref="GI:2624258"
 /db_xref="SWISS-PROT:P26345"
 /translation="MTQTPDRKEGLAVAAQIKTSYKGSVMALGDEAQPISVPTG SILDVALIGIGIPRGRVIEIKPESGGKTTVALHVAANAQAAGVAATPDEHADP DYAKKLGVDTSLLVSQPDTEQALIEIADMLIRSGALDIVIDSAVALPRLHEGEM GSHVGIQARLMSQALRKMTGALNNGTALIFINQLRDKIVGFGSPETTTGKALKE YASVRMDVQREVTLEKIDGTNAVGNRIKRVKVKKNKCLAEGRIEPVTGTHREDVYK RPIHYVAAKDCGTLEHARPVSWFDGTDRVIGLRAGAVIATPDPHVLVEYGNRA AGELRKGRDRAQPRRFGDSAPADAHARLIGYLIIGDRCMGCKGKPIINFORA ALIDVPTRIATIGCAHPQGRISIAIAHRPGRNVAIDCOAGIYGLHAMEKITPN WFEPTIADIVGNLGLFESDQWVSREDTGLRVGYTTTSEQLAHQHLMLRFGV GSVTRVDPTQKRPSIVNGRIQSKRQVEVRISGMDNTAFAESVPMQPRGAALIQ AIBEATQGRRRQSATYLAEMTDVLANLDERGVTAQAAAMIGVASDPRGKMOV LGASRLRRDVRQALADLDKFLHMDIAELRYSVIREVLPTRARTPFLVEELHTL VAGGVVHNSPPEKQAEFDILYKGISREGSLIDMGVOGILRKSGAETPTEGEGLS OGGENARNPLVEADVADAEKIKIRKIKIGLVANVTDDPSNDGYLPAPVP"
 complement(790. .2112)
 /gene="recA"
 /note="recA intein"
 complement(790. .807)
 /gene="recA"
 complement(2197. .2223)
 /note="PS00881 Protein splicing signature"
 complement(2197. .2223)
 /gene="recA"
 complement(2644. .2667)
 /note="PS00321 recA signature"
 complement(2644. .2667)
 /gene="recA"
 /note="PS00017 ATP/GTP-binding site motif A (P-loop)"
 complement(2873. .2877)
 /note="possible RBS upstream of Rv2737c"
 complement(3247. .3453)
 /gene="Rv2738c"
 complement(3247. .3453)
 /gene="Rv2738c"
 complement(3247. .3453)
 /note="Rv2738c, (MTV002.03c), len: 68 aa. Unknown, N-terminus is highly similar to the N-terminus of the

upstream OREFMTV002.07c (78.4% identity in 37 aa overlap); also similar to AL020958|SC4H8_5 Streptomyces coelicolor cosmid 4H8 (64aa) opt: 185 z-score: 283.5 E(): 2.9e-08; 35.7% identity in 63 aa overlap"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein Rv2738c"
 /protein_id="CA15534.1"
 /db_xref="GI:2624260"
 /db_xref="SPTREMBL:O33281"
 /translation="MAGVRLTEFHERVALHFGAAYGSSVLLDHYLTGFDGSAQAQI EDGVEPRDVMRALCADFVPHDRW"
 complement(3464. .4630)
 /gene="Rv2739c"
 complement(3464. .4630)
 /note="Rv2739c, (MTV002.04c), len: 388 aa. Probable transferase, similar to eg. TR:Q51560 (EMBL:L28170) Pseudomonas aeruginosa thymosyl transferase (426 aa), fasta scores: opt: 178 z-score: 226.3 E(): 1.9e-05, 25.9% identity in 425 aa overlap. Equivalent to Mycobacterium leprae protein MLCB33.02c (392 aa); fasta scores: opt: 294.723|MLCB33.2 Mycobacterium leprae cosmid B33 opt: 211.2 z-score: 2364.5 E(): 0; 80.9% identity in 388 aa overlap"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein Rv2739c"
 /protein_id="CA15535.1"
 /db_xref="GI:2624261"
 /db_xref="SPTREMBL:O33282"
 /translation="MRAVAVAGPDGSHSPALALCOFRRAADPTLTGVEVLEAHR AAGIDAVELDGLAATDRDLADGARIHRAAAMAVLNPRALPELVASVITACGG MAALLGIIPWELNPHPLPLPSKGLPAPGASLAGTGRGLRATRKALGRSMRAG LROBAVAVEIGLPPARDPGLRLIATIPALEVERDPPAAVAVVGFHFEPDRVLA IPAGTGPVAVVAPFATGTAGTIGTVALOSITLPETVSGSRVSVSLGADLVPPM AVAGIGSOAELLTRADIVYICGGCHGMVAKTILAVYPMVYVGGGDQNEIANRVROGS AVLIRPLTADALVAANVNEVLSPPREARRRAASVAGADPVVCHDALAAG"
 4674. .5123
 /gene="Rv2740"
 4674. .5123
 /note="Rv2740, (MTV002.05), len: 149 aa. unknown, similar to M. leprae hypothetical protein TR:Q49650 (EMBL:294723) MLCB33.03 (82235-83140) (178 aa), fasta scores: opt: 498 z-score: 538.9 E(): 7.3e-23, 51.6% identity in 161 aa overlap"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein Rv2740"
 /protein_id="CA15536.1"
 /db_xref="GI:2624262"
 /db_xref="SPTREMBL:O33283"
 /translation="MAELTSPETPETTEAIRAVEAFNLQNEDEFDYDAALGDDL VYENVGFSIRIGGRATLILRRMGKRVFEYKIRIGDGAATVTEEDLIIIGLPLV OEWVGVEVDGDRITLMRDYEDVYDMKGLRLQVLAALVYSLKATL"
 5355. .6932
 /gene="PE_PGRS"
 5355. .6932
 /note="Rv2741, (MTV002.06), len: 525 aa. Member of M. tuberculosis PE_PGRS subfamily, similar to many eg.

Query Match 100.0% Score 1237 DB 1: Length 56414:
 Best Local Similarity 100.0% Pred. No. 9.9e-171:
 Matches 1237: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTTGACATTAATGCAACTTCTTCATAGCAGCGTACGATCGAGGGGTAATC 60
 |||||||
 DB 38201 ACTTGCAATTAATGCAACTTCTTCATAGCAGCGTACGATCGAGGGGTAATC 38260
 |||||||
 QY 61 ATGCGCGTGGTATTCGACGACCAACCAACGAATTCGGGTGCCATCACC 120
 |||||||

Db 38261 ATGCGCGTGGATATTCGACGACGACCAAAACAAATTCGGGTGGCATCACCCG 38320
 QY 121 GCCGGGCTGGCGAATAACCCGCTGTCGCAATGAGTGTCTATCCAGGAGGTGCCGA 180
 Db 38321 GCCGGGCTGGCGAATAACCCGCTGTCGCAATGAGTGTCTATCCAGGAGGTGCCGA 38380
 QY 181 GAGGCGTGGCTATACCGACGCGGATTCAGAGGCGGACGCGCACTGTGGCGACC 240
 Db 38381 GAGGCGTGGCTATACCGACGCGGATTCAGAGGCGGACGCGCACTGTGGCGACC 38440
 QY 241 GCCGACAGAGTGTGGCGGACGCTGATTTATGCTCAAGGTCAAGAACCGATAGCGCG 300
 Db 38441 GCCGACAGAGTGTGGCGGACGCTGATTTATGCTCAAGGTCAAGAACCGATAGCGCG 38500
 QY 301 GAATAGCGCGGCTGGGACAGGCGGATCTTGTACGTTCTTGTGATTTGGCGGCTCA 360
 Db 38501 GAATAGCGCGGCTGGGACAGGCGGATCTTGTACGTTCTTGTGATTTGGCGGCTCA 38560
 QY 361 CGTGTTCACGAGATGCTGTTGATTCGCGACGACGCTCAATTCCTACGAGACCGTC 420
 Db 38561 CGTGTTCACGAGATGCTGTTGATTCGCGACGACGCTCAATTCCTACGAGACCGTC 38620
 QY 421 CAGACCGCGGACGCGGACGCTACCCCTGCTTGGCCGATAGCGAATCGCCGCTGACTC 480
 Db 38621 CAGACCGCGGACGCGGACGCTACCCCTGCTTGGCCGATAGCGAATCGCCGCTGACTC 38680
 QY 481 GCCGCCAGGTTGGGCTTACGACGCTGATGCGAACCGAAGGCGCGCGGTGCTGATG 540
 Db 38681 GCCGCCAGGTTGGGCTTACGACGCTGATGCGAACCGAAGGCGCGCGGTGCTGATG 38740
 QY 541 GCGGGGCTGCGCGGCTGCAACCGCGGACGCTGCTGATGCGCGCGCGCGCGCGCG 600
 Db 38741 GCGGGGCTGCGCGGCTGCAACCGCGGACGCTGCTGATGCGCGCGCGCGCGCGCG 38800
 QY 601 TCAACGACGCGCGGCTTGGGCAAGGCGGATGGGCGGACGCTTGTAGACATCAAC 660
 Db 38801 TCAACGACGCGCGGCTTGGGCAAGGCGGATGGGCGGACGCTTGTAGACATCAAC 38860
 QY 661 ATGCAAACTTCGCACTGCAACGCGGAGTTCTGCGCGGATCCACACTGCTACTCA 720
 Db 38861 ATGCAAACTTCGCACTGCAACGCGGAGTTCTGCGCGGATCCACACTGCTACTCA 38920
 QY 721 TCGGCTTACGAGCTGAGAGGTGCGCTCAAAAGTGGCGACTGCTATTTGGGCGCTCTG 780
 Db 38921 TCGGCTTACGAGCTGAGAGGTGCGCTCAAAAGTGGCGACTGCTATTTGGGCGCTCTG 38980
 QY 781 GTGCGAGGCGGACGACCAATATTAGTCTGCAATTCCTGTCGCGATATGAACCA 840
 Db 38981 GTGCGAGGCGGACGACCAATATTAGTCTGCAATTCCTGTCGCGATATGAACCA 39040
 QY 841 GGTGCGGCTGCTGATATATGCAATGACGACGAGGCGGCTGTTGCAAGGCTACAGCG 900
 Db 39041 GGTGCGGCTGCTGATATATGCAATGACGACGAGGCGGCTGTTGCAAGGCTACAGCG 39100
 QY 901 ACCACTTACGACACCGGAGCTTGGCGGACGAGACGCTGTTTACTGCTGGGGAAC 960
 Db 39101 ACCACTTACGACACCGGAGCTTGGCGGACGAGACGCTGTTTACTGCTGGGGAAC 39160
 QY 961 ATGCGCGGCTGCTGCGGACGAGCTGACGCTGACGCGCTGACCAAGCGAGATGCGGTAT 1020
 Db 39161 ATGCGCGGCTGCTGCGGACGAGCTGACGCTGACGCGCTGACCAAGCGAGATGCGGTAT 39220
 QY 1021 ATGCGCGGCTGCTGCGGACGAGCTGACGCTGACGCGCTGACCAAGCGAGATGCGGTAT 1140
 Db 39221 ATGCGCGGCTGCTGCGGACGAGCTGACGCTGACGCGCTGACCAAGCGAGATGCGGTAT 39280
 QY 1081 AAGAGCTTTCACGACGAGAGGCGGCTTACTGTCGAAAGGCTGCGCACTGCGG 1140
 Db 39281 AAGAGCTTTCACGACGAGAGGCGGCTTACTGTCGAAAGGCTGCGCACTGCGG 39340
 QY 1141 GTGCGGCTTACGAGCGCGGACGAGCTGCTGCGCTGACTTCGCGGCTGTTACGCGGAG 1200
 Db 39341 GTGCGGCTTACGAGCGCGGACGAGCTGCTGCGCTGACTTCGCGGCTGTTACGCGGAG 39400

QY 1201 CACAGCTGGGAGTAAGGAGACGATGATGTCGCGG 1237
 Db 39401 CACAGCTGGGAGTAAGGAGACGATGATGTCGCGG 39437
 RESULT 4
 A87613
 LOCUS A87613 1236 bp DNA
 DEFINITION Sequence 10 from Patent WO9836089.
 ACCESSION A87613
 VERSION A87613.1 GI:6736253
 KEYWORDS
 SOURCE
 ORGANISM
 unclassified.
 unclassified.
 unclassified.
 1 (bases 1 to 1236)
 REFERENCES
 FLOHE, L. and Singh, M.
 TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
 Patent: WO 9836089-A-20-AUG-1998
 JOURNAL FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 Location/Qualifiers
 source
 1..1236
 /organism="unclassified"
 /db_xref="taxon:32644"
 BASE COUNT 236 a 395 c 385 g 220 t
 ORIGIN
 Query Match 99.8%; Score 1234.4; DB 5; Length 1236;
 Best Local Similarity 99.9%; Pred. No. 5,4e-170;
 Matches 1235; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATCTGCAATTAATGCACTTTCTCTACTGAAAGCTACAGTATCGAGGAGTATC 60
 Db 1 ATCTTTCAGATTATGCACTTTCTCTACTGAAAGCTACAGTATCGAGGAGTATC 60
 QY 61 ATGCGJGTGCTATTCGACGAGACGACAAACAAACGAAATTCGCGGTGGCATCCCG 120
 Db 61 ATGCGGCTGCTATTCGACGAGACGACAAACAAACGAAATTCGCGGTGGCATCCCG 120
 QY 121 GCCGCGTGGGGAATTAACCGCTGCTGCGCATGAGTGTCTATCAGAGGTGCCGA 180
 Db 121 GCCGCGTGGGGAATTAACCGCTGCTGCGCATGAGTGTCTATCAGAGGTGCCGA 180
 QY 181 GAGGCTGCGGATACGACGAGCGGATTCGAAAGGCGGACGCGCAATCTGCGGACC 240
 Db 181 GAGGCTGCGGATACGACGAGCGGATTCGAAAGGCGGACGCGCAATCTGCGGACC 240
 QY 241 GCCGACAGGCTGCGGACGACGCTGATTTATTCAGAGTCAAGAACCGATAGCGCG 300
 Db 241 GCCGACAGGCTGCGGACGACGCTGATTTATTCAGAGTCAAGAACCGATAGCGCG 300
 QY 301 GAATAGCGGCTGCGGACGAGCTGCTTGTTCACGTTCTTGATTTGGCGGCTCA 360
 Db 301 GAATAGCGGCTGCGGACGAGCTGCTTGTTCACGTTCTTGATTTGGCGGCTCA 360
 QY 361 CGTGTTCACGAGATGCTGTTGATTCGCGGACGAGCTGATTTGCTTACGAGACCGTC 420
 Db 361 CGTGTTCACGAGATGCTGTTGATTCGCGGACGAGCTGATTTGCTTACGAGACCGTC 420
 QY 421 CAGACCGCGGACGCGGACGCTACCCCTGCTTGGCCGATAGCGAATCGCCGCTGACTC 480
 Db 421 CAGACCGCGGACGCGGACGCTACCCCTGCTTGGCCGATAGCGAATCGCCGCTGACTC 480
 QY 481 GCCGCCAGGTTGGCGCTTACCACTGATGCGAACCAAGGCGGCGCGGTGCTGATG 540
 Db 481 GCCGCCAGGTTGGCGCTTACCACTGATGCGAACCAAGGCGGCGCGGTGCTGATG 540
 QY 541 GCGGGGCTGCGCGGCTGCAACCGGCGGACGCTGCTGATGCGCGCGCGGCGCGCGG 600
 Db 541 GCGGGGCTGCGCGGCTGCAACCGGCGGACGCTGCTGATGCGCGCGCGGCGCGCGG 600

QY 601 TACAGCAGCCGCGATGCGCAAGCGGATGGGCGGACCGTTACGGTTCTAGACATCAAC 660
 Db 601 TACAGCAGCCGCGATGCGCAAGCGGATGGGCGGACCGTTACGGTTCTAGACATCAAC 660
 QY 661 ATGACAACTTGGGCAACTGACGCGCGATTTCTGGGCGCGGATTCACACTCTGCTACTCA 720
 Db 661 ATGACAACTTGGGCAACTGACGCGCGATTTCTGGGCGCGGATTCACACTCTGCTACTCA 720
 QY 721 TCGGCTACGAGCTGAGAGGAGGCGGTCGAAAGCTGCGGACCTGATGATGGGCGGCTCTG 780
 Db 721 TCGGCTACGAGCTGAGAGGAGGCGGTCGAAAGCTGCGGACCTGATGATGGGCGGCTCTG 780
 QY 781 GTGCGAGCGCGCAAGGACCGCAATTAATGCTGCAATTCCTGCTGCGGATATGAAACCA 840
 Db 781 GTGCGAGCGCGCAAGGACCGCAATTAATGCTGCAATTCCTGCTGCGGATATGAAACCA 840
 QY 841 GGTGCGGCTGAGTGTGATATAGCCATGACAGGCGGCGTGTTCGAAAGGCTCACAGCCG 900
 Db 841 GGTGCGGCTGAGTGTGATATAGCCATGACAGGCGGCGTGTTCGAAAGGCTCACAGCCG 900
 QY 901 ACCACCTACGACACCGACGCTGCGCGTGCAGACGACGCTGTTTACTGCGTGGCGAAC 960
 Db 901 ACCACCTACGACACCGACGCTGCGCGTGCAGACGACGCTGTTTACTGCGTGGCGAAC 960
 QY 961 ATGCGCGCTGCGTGGCGGACGCTGACCTTACGCGCTGACGACGCGGATGCGCTAT 1020
 Db 961 ATGCGCGCTGCGTGGCGGACGCTGACCTTACGCGCTGACGACGCGGATGCGCTAT 1020
 QY 1021 GTGCTGAGCTGCGGACGATGCTGCGGCGGCGGCGTGCAGGCTGCAATCCGACTACCC 1080
 Db 1021 GTGCTGAGCTGCGGACGATGCTGCGGCGGCGGCGTGCAGGCTGCAATCCGACTACCC 1080
 QY 1081 AAAGGCTTTTGCAGCAGCAGAGGCGGCTTACTGTCCGAACGGGTGCGCACTGCGGG 1140
 Db 1081 AAAGGCTTTTGCAGCAGCAGAGGCGGCTTACTGTCCGAACGGGTGCGCACTGCGGG 1140
 QY 1141 GTGCGGCTTACGAGCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
 Db 1141 GTGCGGCTTACGAGCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
 QY 1201 CACAGCTGCGGAGTAAGGAGGAGGATGATGTCGCC 1236
 Db 1201 CACAGCTGCGGAGTAAGGAGGAGGATGATGTCGCC 1236
 RESULT 5
 A89753 1236 bp DNA PAT 22-JAN-2000
 LOCUS A89753 Sequence 10 from Patent WO9832862.
 DEFINITION A89753
 ACCESSION A89753.1 GI:6738287
 VERSION A89753.1 GI:6738287
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1236)
 AUTHORS Flohe, L. and Singh, M.
 TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
 JOURNAL Patent: WO 9832862-A 30-JUL-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source 1..1236
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 236 a 395 c 385 g 220 t
 ORIGIN
 Query Match 99.8%; Score 1234.4; DB 5; Length 1236;
 Best Local Similarity 99.9%; Pred. No. 5.4e-170;
 Matches 1235; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATCTTCAGATTAATGCAACTTCTTCACTGACGCTACGATGATGAGAGGGGTAATC 60

Db 1 ATCTTCAGATTAATGCAACTTCTTCACTGACGCTACGATGATGAGAGGGGTAATC 60
 QY 61 ATGCGGCTGCTATTCGAGCGGAGGACCAAAACAGATTCGGGTGCGCAATCCCG 120
 Db 61 ATGCGGCTGCTATTCGAGCGGAGGACCAAAACAGATTCGGGTGCGCAATCCCG 120
 QY 121 GCCGGGCTGCGGAACTAACCCGCTGCGCATGAGTGTCTATCCAGGAGGTGCCGA 180
 Db 121 GCCGGGCTGCGGAACTAACCCGCTGCGCATGAGTGTCTATCCAGGAGGTGCCGA 180
 QY 181 GAGGCTGCGGCTATCACCGAGCGGATTTCAAGCGGCGAGCGCGCAATGCTGCGCAC 240
 Db 181 GAGGCTGCGGCTATCACCGAGCGGATTTCAAGCGGCGAGCGCGCAATGCTGCGCAC 240
 QY 241 GCCGACGAGTGTGGGCGGCGGCTGATTTATGCTCAAGGCTAAAGACGATAGCGCG 300
 Db 241 GCCGACGAGTGTGGGCGGCGGCTGATTTATGCTCAAGGCTAAAGACGATAGCGCG 300
 QY 301 GAATACGCGCGCTGCGGACGAGCGGAGATCTTGTACGTTCTTGCATTTGGCGGCTCA 360
 Db 301 GAATACGCGCGCTGCGGACGAGCGGAGATCTTGTACGTTCTTGCATTTGGCGGCTCA 360
 QY 361 CGTCTTGCAGCGGATGCTTGTGATTCGCGGACGACGCTCAATGCTTACGAGACGCTC 420
 Db 361 CGTCTTGCAGCGGATGCTTGTGATTCGCGGACGACGCTCAATGCTTACGAGACGCTC 420
 QY 421 CAGACCGCGGAGCGGACGCTACCTGCTTGTCCCGATGAGCGAAGTTCGCGGCTGACTC 480
 Db 421 CAGACCGCGGAGCGGACGCTACCTGCTTGTCCCGATGAGCGAAGTTCGCGGCTGACTC 480
 QY 481 GCGCGCGAGTGGCGGCTTACACCTGATGCGAACCAGGGGCGGCGGTGTGCTGATG 540
 Db 481 GCGCGCGAGTGGCGGCTTACACCTGATGCGAACCAGGGGCGGCGGTGTGCTGATG 540
 QY 541 GCGCGGCTGCGGCGGCTTGAACCGGCGGCGGCTGCTGATGCGGCGGCGGCGGCGG 600
 Db 541 GCGCGGCTGCGGCGGCTTGAACCGGCGGCGGCTGCTGATGCGGCGGCGGCGGCGG 600
 QY 601 TACAGCAGCGCGGACGCTGCGCAAGGAGTGGGCGGCGGCTTACGCTTACGATCAAC 660
 Db 601 TACAGCAGCGCGGACGCTGCGCAAGGAGTGGGCGGCGGCTTACGCTTACGATCAAC 660
 QY 661 ATGACAACTTGGGCAACTGACGCGCGATTTGCGGCGGAGATTCACACTCTGCTACTCA 720
 Db 661 ATGACAACTTGGGCAACTGACGCGCGATTTGCGGCGGAGATTCACACTCTGCTACTCA 720
 QY 721 TCGGCTACGAGCTGAGAGGAGGCGGTCGAAAGCTGCGGACCTGATGATGGGCGGCTCTG 780
 Db 721 TCGGCTACGAGCTGAGAGGAGGCGGTCGAAAGCTGCGGACCTGATGATGGGCGGCTCTG 780
 QY 781 GTGCGAGCGCGCAAGGACCGCAATTAATGCTGCAATTCCTGCTGCGGATATGAAACCA 840
 Db 781 GTGCGAGCGCGCAAGGACCGCAATTAATGCTGCAATTCCTGCTGCGGATATGAAACCA 840
 QY 841 GGTGCGGCTGAGTGTGATATAGCCATGACAGGCGGCGTGTTCGAAAGGCTCACAGCCG 900
 Db 841 GGTGCGGCTGAGTGTGATATAGCCATGACAGGCGGCGTGTTCGAAAGGCTCACAGCCG 900
 QY 901 ACCACCTACGACACCGACGCTGCGCGTGCAGACGACGCTGTTTACTGCGTGGCGAAC 960
 Db 901 ACCACCTACGACACCGACGCTGCGCGTGCAGACGACGCTGTTTACTGCGTGGCGAAC 960
 QY 961 ATGCGCGCTGCGTGGCGGACGCTGACCTTACGCGCTGACGACGCGGATGCGCTAT 1020
 Db 961 ATGCGCGCTGCGTGGCGGACGCTGACCTTACGCGCTGACGACGCGGATGCGCTAT 1020
 QY 1021 GTGCTGAGCTGCGGACGATGCTGCGGCGGCGGCGTGCAGGCTGCAATCCGACTACCC 1080
 Db 1021 GTGCTGAGCTGCGGACGATGCTGCGGCGGCGGCGTGCAGGCTGCAATCCGACTACCC 1080
 QY 1081 AAAGGCTTTTGCAGCAGCAGAGGCGGCTTACTGTCCGAACGGGTGCGCACTGCGGG 1140

Db 1081 AAGGTCCTTTCAGACGACGAGAGGGGCGTTACTCTCCGAACGGGTGGCCACGACCTGGGG 1140

QY 1141 GTGCGTTACGAGCCGCGCAGCGTGTGGCTGACTCTCGGCGCTGTACCGCGAG 1200

Db 1141 GTGCGTTACGAGCCGCGCAGCGTGTGGCTGACTCTCGGCGCTGTACCGCGAG 1200

QY 1201 CACACGTGGAGTAAGGAGGAGATGATCGGCC 1236

Db 1201 CACACGTGGAGTAAGGAGGAGATGATCGGCC 1236

RESULT 6

LOCUS A87606 1235 bp DNA PAT 22-JAN-2000

DEFINITION Sequence 3 from Patent WO9836089.

ACCESSION A87606

VERSION A87606.1 GI:6736246

KEYWORDS

SOURCE

ORGANISM

REFERENCE 1 (bases 1 to 1235)

AUTHORS Flohe, L. and Singh, M.

TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE

JOURNAL Patent: WO 9836089-A 20-AUG-1998;

FEATURES

source 1..1235

location/Qualifiers

1..1235

/organism="unidentified"

/db_xref="taxon:32644"

BASE COUNT 236 a 394 c 382 g 220 t 3 others

ORIGIN

Query Match 99.5%; Score 1230.4; DB 5; Length 1235;

Best Local Similarity 99.7%; Pred. No. 2.1e-169;

Matches 1231; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATCTTCAGATTATGAACTTCTCATCTGACGATGACGATGACGAGGGTAAATC 60

Db 1 ATCTTCAGATTATGAACTTCTCATCTGACGATGACGATGACGAGGGTAAATC 60

QY 61 ATGCGGCGTGGTATTCGACGACGACCAAAACAGCAATTCGGGTGGCCATCACCCG 120

Db 61 ATGCGGCGTGGTATTCGACGACGACCAAAACAGCAATTCGGGTGGCCATCACCCG 120

QY 121 GCCGGCGTGGGACCTACCCCTGCTGCGCATGAGGTGCTCATCCAGGAGGTGCCGA 180

Db 121 GCCGGCGTGGGACCTACCCCTGCTGCGCATGAGGTGCTCATCCAGGAGGTGCCGA 180

QY 181 GAGGGGTGGCTATCAACCGGCGGAGATTCAAGGCGGAGCGCGCACTGGTGGGAC 240

Db 181 GAGGGGTGGCTATCAACCGGCGGAGATTCAAGGCGGAGCGCGCACTGGTGGGAC 240

QY 241 GCGGACGAGTGTGGCGCGACGCTGATTTATTCCTCAAGGTCAAAAGACGATAGCGCG 300

Db 241 GCGGACGAGTGTGGCGCGACGCTGATTTATTCCTCAAGGTCAAAAGACGATAGCGCG 300

QY 301 GATTAAGGCGCGCTGGGACGACGCGGAGATTTTTCACGTTCTTGATTTGGCCGCTCA 360

Db 301 GATTAAGGCGCGCTGGGACGACGCGGAGATTTTTCACGTTCTTGATTTGGCCGCTCA 360

QY 361 CGTGTTCGACGATGCTGTGATTCGCGACACGATTAATGCTTACGAGACCTGTC 420

Db 361 CGTGTTCGACGATGCTGTGATTCGCGACACGATTAATGCTTACGAGACCTGTC 420

QY 421 CAGACGCGGAGCGGCGACTACCCCTGCTTCCCGCATGAGCGAAGTGGCGGTGCACTC 480

Db 421 CAGACGCGGAGCGGCGACTACCCCTGCTTCCCGCATGAGCGAAGTGGCGGTGCACTC 480

QY 481 GCGGCGGAGTTGGCGCTTACCAACGATGATGCAACCAAGGGGCGCGGTGTGCTGATG 540

Db 481 GCGGCGGAGTTGGCGCTTACCAACGATGATGCAACCAAGGGGCGCGGTGTGCTGATG 540

QY 541 GGGGGGTCGCGCGGCTGCAACCGCGCGACGTCGTGTGATGGCGCGCGACCGCGCG 600

Db 541 GGGGGGTCGCGCGGCTGCAACCGCGCGACGTCGTGTGATGGCGCGCGACCGCGCG 600

QY 601 TACAAGCGACCGCGCATGCCCAAGCGCATGGCGGACGCTTACGGTTTACAATCAAC 660

Db 601 TACAAGCGACCGCGCATGCCCAAGCGCATGGCGGACGCTTACGGTTTACAATCAAC 660

QY 661 ATGACAACTTCGCGCACTGACGCGCGAGTCTGCGGCGGATCCACACTGCTACTCA 720

Db 661 ATGACAACTTCGCGCACTGACGCGCGAGTCTGCGGCGGATCCACACTGCTACTCA 720

QY 721 TCGGCTACGAGCTGAGGCTGCCGTCACCAACGTCGCGGACCTGATTTGGGCGCTCTG 780

Db 721 TCGGCTACGAGCTGAGGCTGCCGTCACCAACGTCGCGGACCTGATTTGGGCGCTCTG 780

QY 781 GTGCGAGCGCGCAAGGACCCCAATTAATGCTGCAATTCATTTGCGCGATTAAGAACCA 840

Db 781 GTGCGAGCGCGCAAGGACCCCAATTAATGCTGCAATTCATTTGCGCGATTAAGAACCA 840

QY 841 GGTGGGTACTGTGTGATTAACCATTCGACACGAGGCGGCTGTTGGAAGGCTACGACCG 900

Db 841 GGTGGGTACTGTGTGATTAACCATTCGACACGAGGCGGCTGTTGGAAGGCTACGACCG 900

QY 901 ACCACTAGACACCCGACGTCGCGCGCTGCGACGACGACCTGTTTACGCTGGCGAAC 960

Db 901 ACCACTAGACACCCGACGTCGCGCGCTGCGACGACGACCTGTTTACGCTGGCGAAC 960

QY 961 ATGCCCGCTCGGTGCGCAAGACGTCGACCTACGCGCTGACCAACGCGACGATGCGGAT 1020

Db 961 ATGCCCGCTCGGTGCGCAAGACGTCGACCTACGCGCTGACCAACGCGACGATGCGGAT 1020

QY 1021 GTGCTGAGCTTCCGACCATGCTGGCGGGCGGCTGCGGCTGCAATCCGCGACTAGCC 1080

Db 1021 GTGCTGAGCTTCCGACCATGCTGGCGGGCGGCTGCGGCTGCAATCCGCGACTAGCC 1080

QY 1081 AAGGCTTTCGACGACGACGAGGCGGCTTACTGTCGCAAGCGCTGGCAGCGACTGGGG 1140

Db 1081 AAGGCTTTCGACGACGACGAGGCGGCTTACTGTCGCAAGCGCTGGCAGCGACTGGGG 1140

QY 1141 GTGCGTTACCGAGCGCGCCGACGCTGCTGCGCTGACTTCGCGCGCTGTACCGCGAG 1200

Db 1141 GTGCGTTACCGAGCGCGCCGACGCTGCTGCGCTGACTTCGCGCGCTGTACCGCGAG 1200

QY 1201 CACACGTGGAGTAAGGAGGACGATGTCGCC 1235

Db 1201 CACACGTGGAGTAAGGAGGACGATGTCGNC 1235

RESULT 7

LOCUS A89746 1235 bp DNA PAT 22-JAN-2000

DEFINITION Sequence 3 from Patent WO9832862.

ACCESSION A89746

VERSION A89746.1 GI:6738280

KEYWORDS

SOURCE

ORGANISM

REFERENCE 1 (bases 1 to 1235)

AUTHORS Flohe, L. and Singh, M.

TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM

JOURNAL Patent: WO 9832862-A 30-JUL-1998;

FEATURES

source 1..1235

location/Qualifiers

1..1235

/organism="unidentified"

/db_xref="taxon:32644"

BASE COUNT 236 a 394 c 382 g 220 t 3 others

ORIGIN

Query Match 99.5%; Score 1230.4; DB 5; Length 1235;
 Best Local Similarity 99.7%; Pred. No. 2,1e-169;
 Matches 1231; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 ATCTTGCAATTAATGAACTTCTCTCATGAAAGCTAGACGATGAGAGGGGTAATC 60
   |||||||
Db 1 ATCTTGCAATTAATGAACTTCTCTCATGAAAGCTAGACGATGAGAGGGGTAATC 60
QY 61 ATGCGGCTGGTATTCAGACGAGACCAAAACAAAGAAATTCGGGTGGCCATCAACCCG 120
   |||||||
Db 61 ATGCGGCTGGTATTCAGACGAGACCAAAACAAAGAAATTCGGGTGGCCATCAACCCG 120
QY 121 GCGGCGTGGGAAATAACCGCTGCTGGCATGAGTCTCATCAGGACAGTGGCGGA 180
   |||||||
Db 121 GCGGCGTGGGAAATAACCGCTGCTGGCATGAGTCTCATCAGGACAGTGGCGGA 180
QY 181 GAGGCTCGGCTATCACGACGAGGATTTCAAGGCGGACGCGCACTGGTGGCAC 240
   |||||||
Db 181 GAGGCTCGGCTATCACGACGAGGATTTCAAGGCGGACGCGCACTGGTGGCAC 240
QY 241 GCGGACAGTGTGGGCGGACGCTGATTTATGCTCAAGAGTCAAGAACGATAGCGCG 300
   |||||||
Db 241 GCGGACAGTGTGGGCGGACGCTGATTTATGCTCAAGAGTCAAGAACGATAGCGCG 300
QY 301 GAATACGCGCGCTGCGACACGCGGAGATCTTGTCAGCTCTTGATTTGGCCGCTCA 360
   |||||||
Db 301 GAATACGCGCGCTGCGACACGCGGAGATCTTGTCAGCTCTTGATTTGGCCGCTCA 360
QY 361 CGTCTGACAGGATGCGTTGTTGATTCGGGACACGCTCAATGCTTACAGACCGTC 420
   |||||||
Db 361 CGTCTGACAGGATGCGTTGTTGATTCGGGACACGCTCAATGCTTACAGACCGTC 420
QY 421 CAGACCGCGGACGCGGACATACCTGCTTGGCCGATGAGGAAAGTGGCGGTGAC 480
   |||||||
Db 421 CAGACCGCGGACGCGGACATACCTGCTTGGCCGATGAGGAAAGTGGCGGTGAC 480
QY 481 GCGGCGGCTGGGCTTACCACTGATGCAACCCAAAGGCGCGGCTGCTGATG 540
   |||||||
Db 481 GCGGCGGCTGGGCTTACCACTGATGCAACCCAAAGGCGCGGCTGCTGATG 540
QY 541 GCGGCGGCTGGGCTGCAACGCGGCGGCTGCTGATGAGGCGGCGGCGCGCG 600
   |||||||
Db 541 GCGGCGGCTGGGCTGCAACGCGGCGGCTGCTGATGAGGCGGCGGCGCGCGCG 600
QY 601 TACAAACGAGCGCGCATGCGCAAGGCGATGGCGGCGGCTTACGTTCTAGACATCAAC 660
   |||||||
Db 601 TACAAACGAGCGCGCATGCGCAAGGCGATGGCGGCGGCTTACGTTCTAGACATCAAC 660
QY 661 ATGCAACGAGCGCGCATGCGCAAGGCGATGGCGGCGGCTTACGTTCTAGACATCAAC 720
   |||||||
Db 661 ATGCAACGAGCGCGCATGCGCAAGGCGATGGCGGCGGCTTACGTTCTAGACATCAAC 720
QY 721 TCGGCTACGAGCTGAGGCTGCGTCAAAAGCTGCGGACCTGATTTGGGCGCTCTG 780
   |||||||
Db 721 TCGGCTACGAGCTGAGGCTGCGTCAAAAGCTGCGGACCTGATTTGGGCGCTCTG 780
QY 781 GTCGAGGCGGCGCAAGGCGCAAAATTAATGCTGCAATTTGCTGCGCATATGAACCA 840
   |||||||
Db 781 GTCGAGGCGGCGCAAGGCGCAAAATTAATGCTGCAATTTGCTGCGCATATGAACCA 840
QY 841 GTCGAGGCTGATGATATGATGCAACGAGGCGGCTGTTGGAAGGCTCAGACCG 900
   |||||||
Db 841 GTCGAGGCTGATGATATGATGCAACGAGGCGGCTGTTGGAAGGCTCAGACCG 900
QY 901 ACCACCTACGACCGCGAGCTTGGCTGACGACGAGCGTCTTTACTGGGTGGCGAAC 960
   |||||||
Db 901 ACCACCTACGACCGCGAGCTTGGCTGACGACGAGCGTCTTTACTGGGTGGCGAAC 960
QY 961 ATGCGGCTGGTGGCGGAGGCTGCAAGGCTGAGGCTGAGCAAGGCGATGCGCGTAT 1020
   |||||||
Db 961 ATGCGGCTGGTGGCGGAGGCTGCAAGGCTGAGGCTGAGCAAGGCGATGCGCGTAT 1020
QY 1021 GTGCTGAGCTTGGCGGACCATGCTGCGGCGGCTGCGGCTGCAATCGGCACTACCC 1080

```

Query Match 99.3%; Score 1228; DB 5; Length 1228;
 Best Local Similarity 100.0%; Pred. No. 4,6e-169;
 Matches 1228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATCTTGCAATTAATGAACTTCTCTCATGAAAGCTAGACGATGAGAGGGGTAATC 60
   |||||||
Db 1 ATCTTGCAATTAATGAACTTCTCTCATGAAAGCTAGACGATGAGAGGGGTAATC 60
QY 61 ATGCGGCTGGTATTCAGACGAGACCAAAACAAAGAAATTCGGGTGGCCATCAACCCG 120
   |||||||
Db 61 ATGCGGCTGGTATTCAGACGAGACCAAAACAAAGAAATTCGGGTGGCCATCAACCCG 120
QY 121 GCGGCGTGGGAAATAACCGCTGCTGGCATGAGTCTCATCAGGACAGTGGCGGA 180
   |||||||
Db 121 GCGGCGTGGGAAATAACCGCTGCTGGCATGAGTCTCATCAGGACAGTGGCGGA 180
QY 181 GAGGCTCGGCTATCACGAGCGGATTTCAAGGCGGACGCGCAACTGTCGCGCAC 240
   |||||||
Db 181 GAGGCTCGGCTATCACGAGCGGATTTCAAGGCGGACGCGCAACTGTCGCGCAC 240
QY 241 GCGGACAGTGTGGGCGGACGCTGATTTATGCTCAAGGTCAAAGAACGATAGCGCG 300
   |||||||
Db 241 GCGGACAGTGTGGGCGGACGCTGATTTATGCTCAAGGTCAAAGAACGATAGCGCG 300
QY 301 GAATACGCGCGCTGCGACACGCGGAGATCTTGTCAGCTCTTGATTTGGCCGCTCA 360
   |||||||
Db 301 GAATACGCGCGCTGCGACACGCGGAGATCTTGTCAGCTCTTGATTTGGCCGCTCA 360
QY 361 CGTCTGACAGGATGCGTTGTTGATTCGGGACACGCTCAATGCTTACGAGACCGTC 420
   |||||||
Db 361 CGTCTGACAGGATGCGTTGTTGATTCGGGACACGCTCAATGCTTACGAGACCGTC 420
QY 421 CAGACCGCGGACGCGGACATACCTGCTTGGCCGATGAGGAAAGTGGCGGTGAC 480

```

Db 421 CAGACGGCCGACGGGCACTACCCCTGCTGCCCCGATGAGGAAAGTGCCTGCTGACTC 480
 QY 481 GCCGCCAGAGTTGGGCTTACCACTGATGAGAACCAAGAGGGGGGGGCTGTGATG 540
 Db 481 GCCGCCAGAGTTGGGCTTACCACTGATGAGAACCAAGAGGGGGGGGCTGTGATG 540
 QY 541 GCGGGGGTGGCGGCTGAGAACCGGCGCACTGTGTGTATGCGCGCGGACCGCGGC 600
 Db 541 GCGGGGGTGGCGGCTGAGAACCGGCGCACTGTGTGTATGCGCGCGGACCGCGGC 600
 QY 601 TACAAAGAGCGCGCATGCGCAAGGAGGATGGGCGGACGGTTAGGTTCTAGATCAAC 660
 Db 601 TACAAAGAGCGCGCATGCGCAAGGAGGATGGGCGGACGGTTAGGTTCTAGATCAAC 660
 QY 661 ATGCAAAACTTGGGCACTGACCGCGGATGTTCTGCGCGCGGATCCACACTGCTACTCA 720
 Db 661 ATGCAAAACTTGGGCACTGACCGCGGATGTTCTGCGCGCGGATCCACACTGCTACTCA 720
 QY 721 TCGGCTACGAGCTGAGAGGTGGCTCAAGGTCGCACTGCTGATGGGGCGCTCTG 780
 Db 721 TCGGCTACGAGCTGAGAGGTGGCTCAAGGTCGCACTGCTGATGGGGCGCTCTG 780
 QY 781 GTGCCAGCGCGCAAGGACCCCAATTAAGTCTGCAATTCAGTTGCGCGCATATGAACA 840
 Db 781 GTGCCAGCGCGCAAGGACCCCAATTAAGTCTGCAATTCAGTTGCGCGCATATGAACA 840
 QY 841 GGTGGGTAAGTGTGATATAGCATGACCAAGGCGGCTGTTGAAAGGCTACAGACCG 900
 Db 841 GGTGGGTAAGTGTGATATAGCATGACCAAGGCGGCTGTTGAAAGGCTACAGACCG 900
 QY 901 ACCACCTAGACCAACCGCGCTGCGTGCAGACAGCTGTTTACTGGGTCGCGAAC 960
 Db 901 ACCACCTAGACCAACCGCGCTGCGTGCAGACAGCTGTTTACTGGGTCGCGAAC 960
 QY 961 ATGCCCGCTGCTGCTGCGCAAGAGAGTGCAGCTACGCGCTGACCAAGCGACATGCCGTAT 1020
 Db 961 ATGCCCGCTGCTGCTGCGCAAGAGAGTGCAGCTACGCGCTGACCAAGCGACATGCCGTAT 1020
 QY 1021 GTGCTGAGAGTTGGCGACATGAGTGGGGGGGGGGGGTCCGGTCCAAATCCGGCACTACCC 1080
 Db 1021 GTGCTGAGAGTTGGCGACATGAGTGGGGGGGGGGGGTCCGGTCCAAATCCGGCACTACCC 1080
 QY 1081 AAAGGCTTTTGAGCAGCAAGAGGGGGGCTTACTGTCGAAGGGGTGGCCACGACCTGGGG 1140
 Db 1081 AAAGGCTTTTGAGCAGCAAGAGGGGGGCTTACTGTCGAAGGGGTGGCCACGACCTGGGG 1140
 QY 1141 GTGCCGTTACCGAGCGCGCGCAAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
 Db 1141 GTGCCGTTACCGAGCGCGCGCAAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
 QY 1201 CACAGCTGGGAGTAAGGGAAGCGATGA 1228
 Db 1201 CACAGCTGGGAGTAAGGGAAGCGATGA 1228

RESULT 9

LOCUS A89748 1228 bp DNA
 DEFINITION Sequence 5 from Patent WO9832862.
 ACCESSION A89748
 VERSION A89748.1 GI:6738282
 KEYWORDS
 ORGANISM
 SOURCE
 unclassified.
 unclassified.
 unclassified.

REFERENCE 1 (bases 1 to 1228)
 FIOHE L. and Singh M.

TITILE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARIUM
 JOURNAL Patent: WO 9832862-A 30-JUL-1998;

FEATURES
 source
 1. 1228
 Location/Qualifiers
 /organism="unclassified"

BASE COUNT 236 a 391 c 382 g 219 t
 ORIGIN
 Query Match 99.3%; Score 1228; DB 5; Length 1228;
 Best Local Similarity 100.0%; Pred. No. 4, 6e-169;
 Matches 1228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 ATCTTGAGATTATCCGACCTTCTTATCTAGAGCTACAGATGAGAGGGTATC 60
 Db 1 ATCTTGAGATTATCCGACCTTCTTATCTAGAGCTACAGATGAGAGGGTATC 60
 QY 61 ATGCGCGTGGTATTCGACGAGACCAAAACCAAAATTCGCGGTGACATCACCCG 120
 Db 61 ATGCGCGTGGTATTCGACGAGACCAAAACCAAAATTCGCGGTGACATCACCCG 120
 QY 121 GCGGCGTGGGGAATTAACCCGTGTGCGCATGAGGTGCTCATCGAGGATGCCGA 180
 Db 121 GCGGCGTGGGGAATTAACCCGTGTGCGCATGAGGTGCTCATCGAGGATGCCGA 180
 QY 181 GAGGCTGCGCTATCAGCGAGCGGATTTCAAGGCGGCGAGCGCGCAACTGTGCGCAC 240
 Db 181 GAGGCTGCGCTATCAGCGAGCGGATTTCAAGGCGGCGAGCGCGCAACTGTGCGCAC 240
 QY 241 GCCGACAGGTGTGGCGGAGCTGATTTATTTGCTCAAGGTCAAAAGACGATAGCGCG 300
 Db 241 GCCGACAGGTGTGGCGGAGCTGATTTATTTGCTCAAGGTCAAAAGACGATAGCGCG 300
 QY 301 GAATAGGCGCGCTGCGACACGCGGACATCTGTTCACGTTCTTGCAATTGGCGCGTCA 360
 Db 301 GAATAGGCGCGCTGCGACACGCGGACATCTGTTCACGTTCTTGCAATTGGCGCGTCA 360
 QY 361 CGTCTTGACACGATGCTGTTGATGATTCGCGACACGCTCAATTGCTACGAGACGCTC 420
 Db 361 CGTCTTGACACGATGCTGTTGATGATTCGCGACACGCTCAATTGCTACGAGACGCTC 420
 QY 421 CAGACCGCGGAGCGGCGCACTACCCCTGCTGCGCGCATGAGCGAAGTGCCTGCACTC 480
 Db 421 CAGACCGCGGAGCGGCGCACTACCCCTGCTGCGCGCATGAGCGAAGTGCCTGCACTC 480
 QY 481 GCGCGCCAGGTTGGCGCTTACCACTGATGCAAGCCAAAGGGGGCGGGTGTGATG 540
 Db 481 GCGCGCCAGGTTGGCGCTTACCACTGATGCAAGCCAAAGGGGGCGGGTGTGATG 540
 QY 541 GCGGGGGTGGCGGCTGCAAGCGGCGAGCTGTGTGATTCGCGCGGCGACCGCGGC 600
 Db 541 GCGGGGGTGGCGGCTGCAAGCGGCGAGCTGTGTGATTCGCGCGGCGACCGCGGC 600
 QY 601 TACAAAGAGCGCGCATGCTGCAAGCGGATGGCGCGGAGCTTACGCTTACGATCAAC 660
 Db 601 TACAAAGAGCGCGCATGCTGCAAGCGGATGGCGCGGAGCTTACGCTTACGATCAAC 660
 QY 661 ATGCAAAACTTGGGCACTGACCGCGGATGTTCTGCGCGGATCCACACTGCTACTCA 720
 Db 661 ATGCAAAACTTGGGCACTGACCGCGGATGTTCTGCGCGGATCCACACTGCTACTCA 720
 QY 721 TCGGCTACGAGCTGAGAGGTGGCTCAAGGTCGCACTGCTGATGGGGCGCTCTG 780
 Db 721 TCGGCTACGAGCTGAGAGGTGGCTCAAGGTCGCACTGCTGATGGGGCGCTCTG 780
 QY 781 GTGCCAGCGCGCAAGGACCCCAATTAAGTCTGCAATTCAGTTGCGCGCATATGAACA 840
 Db 781 GTGCCAGCGCGCAAGGACCCCAATTAAGTCTGCAATTCAGTTGCGCGCATATGAACA 840
 QY 841 GGTGGGTAAGTGTGATATAGCATGACCAAGGCGGCTGTTGAAAGGCTACAGACCG 900
 Db 841 GGTGGGTAAGTGTGATATAGCATGACCAAGGCGGCTGTTGAAAGGCTACAGACCG 900
 QY 901 ACCACCTAGACCAACCGCGCTGCGTGCAGACAGCTGTTTACTGGGTCGCGAAC 960
 Db 901 ACCACCTAGACCAACCGCGCTGCGTGCAGACAGCTGTTTACTGGGTCGCGAAC 960

JOURNAL Patent: WO 9832862-A 30-JUL-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
Location/Qualifiers
source 1..1235
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 235 a 395 c 384 g 220 t 1 others
ORIGIN

Query Match 98.7% Score 1221.4; DB 5; Length 1235;
Best Local Similarity 99.8%; Pred. No. 4.1e-168;
Matches 1233; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```

QY 1 ATCTTGAGATTATGAACTTCTTCATCTAGAGCGTACAGATTCAGAGAGGGTAATC 60
DB 1 ATCTTGAGATTATGAACTTCTTCATCTAGAGCGTACAGATTCAGAGAGGGTAATC 60
QY 61 ATGCGCGTCCGATTCGCGAGACCAAAACAAAGAAATTCGGGTGGCCATCACCCCG 120
DB 61 ATGCGCGTCCGATTCGCGAGACCAAAACAAAGAAATTCGGGTGGCCATCACCCCG 120
QY 121 GCCGCGCTGCGGGAATTAACCCGCTGTGGCCATGAGTCTCATCAGAGAGTCCGGA 180
DB 121 GCCGCGCTGCGGGAATTAACCCGCTGTGGCCATGAGTCTCATCAGAGAGTCCGGA 180
QY 181 GAGGCGCTGCGCTATACCGAGCGGATTTCAAGCGCGAGCGCGCAACTGGTGGCACC 240
DB 181 GAGGCGCTGCGCTATACCGAGCGGATTTCAAGCGCGAGCGCGCAACTGGTGGCACC 240
QY 241 GCCGCGAGGTTGGGCGGAGCGCTATTATTCGTCAGAGTCAAGAACCGG AGCGGCG 300
DB 241 GCCGCGAGGTTGGGCGGAGCGCTATTATTCGTCAGAGTCAAGAACCGG AGCGGCG 300
QY 301 GAATACGCGCGCGCTCGACACGCGGAGATCTTGTTCACGTTCTTTCATTTGGCCGCTCA 360
DB 301 GAATACGCGCGCGCTCGACACGCGGAGATCTTGTTCACGTTCTTTCATTTGGCCGCTCA 360
QY 361 CGTCTTGCACCGATGCGCTTGTGATTCGCGACACAGTAAATGCGTACAGACCGCTC 420
DB 361 CGTCTTGCACCGATGCGCTTGTGATTCGCGACACAGTAAATGCGTACAGACCGCTC 420
QY 421 CAGACCGCGAGCGGCGCACTACCCCTGCTGGCCCGATGAGAGAAAGCGCGGTGACTC 480
DB 421 CAGACCGCGAGCGGCGCACTACCCCTGCTGGCCCGATGAGAGAAAGCGCGGTGACTC 480
QY 481 GCCGCGAGGTTGGGCTTACCACTGATGCGAACCCAAAGGGGCGCGGTGCTGATG 540
DB 481 GCCGCGAGGTTGGGCTTACCACTGATGCGAACCCAAAGGGGCGCGGTGCTGATG 540
QY 541 GCCGCGGTCGCGCGGTGGAACCGGCGGACGTCGTGATTCGCGCGCGGCGCGCG 600
DB 541 GCCGCGGTCGCGCGGTGGAACCGGCGGACGTCGTGATTCGCGCGCGGCGCGCG 600
QY 601 TACAACGAGCGCGGATCGCAACGCGATGGGCGGAGCGCTTACGCTCTAGACATCAAC 660
DB 601 TACAACGAGCGCGGATCGCAACGCGATGGGCGGAGCGCTTACGCTCTAGACATCAAC 660
QY 661 ATGCGCAAACTTGGGCACTGAGAGCGGATTCGCGCGCGGATTCAGACCTGATCTCA 720
DB 661 ATGCGCAAACTTGGGCACTGAGAGCGGATTCGCGCGCGGATTCAGACCTGATCTCA 720
QY 721 TCGGCGCTGAGAGTGGGCTTCAAGATGCGGAGTGGGCGGCTGCTG 780
DB 721 TCGGCGCTGAGAGTGGGCTTCAAGATGCGGAGTGGGCGGCTGCTG 780
QY 781 GTGCGAGCGCGCAAGGCGCAAAATTTAGTCTGAATTCATCTTGGCGCAATATAACA 840
DB 781 GTGCGAGCGCGCAAGGCGCAAAATTTAGTCTGAATTCATCTTGGCGCAATATAACA 840
QY 841 GGTGCGGCTACTGATGATATAGCATGAGCGAGGCGGCTTTTGAAGGCTCAGACCG 900
DB 841 GGTGCGGCTACTGATGATATAGCATGAGCGAGGCGGCTTTTGAAGGCTCAGACCG 900

```

```

QY 901 ACCACCTACGACGACCCGAGCTGCGCGTGCACACACAGCTGTTTACTGCTGGGCAAC 960
DB 901 ACCACCTACGACGACCCGAGCTGCGCGTGCACACACAGCTGTTTACTGCTGGGCAAC 960
QY 961 ATGCGCGCTGCTGCGGAGAGCTGACATCCGCTGACCAACGAGAGATCCGTAT 1020
DB 961 ATGCGCGCTGCTGCGGAGAGCTGACATCCGCTGACCAACGAGAGATCCGTAT 1020
QY 1021 GTGCTCAGAGTTCGCGACATGCTGGCGGCGGCTGCGGCTGGAATCCGCGACTAGCC 1080
DB 1021 GTGCTCAGAGTTCGCGACATGCTGGCGGCGGCTGCGGCTGGAATCCGCGACTAGCC 1080
QY 1081 AAAGGTCTTTCGAGCGACGAAGAGGCGCTTACTGTCCAGAGGCGGCGACCGACCTGGGG 1140
DB 1081 AAAGGTCTTTCGAGCGACGAAGAGGCGCTTACTGTCCAGAGGCGGCGACCGACCTGGGG 1140
QY 1141 GTGCGCTTACCGAGCGCGCGCGAGCGTGTGCTGCTGACTCTGCGCGCTGTTAGCGCGAG 1200
DB 1141 GTGCGCTTACCGAGCGCGCGCGAGCGTGTGCTGCTGACTCTGCGCGCTGTTAGCGCGAG 1200
QY 1201 CACAGTCGCGGAGTAAAGGAGGATGATGTCGCGCC 1236
DB 1201 CACAGTCGCGGAGTAAAGGAGGATGATGTCGCGCC 1236

```

RESULT 12

LOCUS A87611 1235 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 8 from Patent WO9836089.
ACCESSION A87611
VERSION A87611.1 GI:6736251
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.

REFERENCE

1 (bases 1 to 1235)
FLOHE/L. and Singh/M.
TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
Patent: WO 9836089-A 20-AUG-1998;
JOURNAL FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)

FEATURES

source

BASE COUNT 236 a 394 c 385 g 220 t
ORIGIN

Query Match 98.7% Score 1220.8; DB 5; Length 1235;
Best Local Similarity 99.8%; Pred. No. 5e-168;
Matches 1233; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```

QY 1 ATCTTGAGATTATGAACTTCTTCATCTAGAGCGTACAGATTCAGAGAGGGTAATC 60
DB 1 ATCTTGAGATTATGAACTTCTTCATCTAGAGCGTACAGATTCAGAGAGGGTAATC 60
QY 61 ATGCGCGTCCGATTCGCGAGACCAAAACAAAGAAATTCGGGTGGCCATCACCCCG 120
DB 61 ATGCGCGTCCGATTCGCGAGACCAAAACAAAGAAATTCGGGTGGCCATCACCCCG 120
QY 121 GCCGCGCTGCGGGAATTAACCCGCTGTGGCCATGAGTCTCATCAGAGAGTCCGGA 180
DB 121 GCCGCGCTGCGGGAATTAACCCGCTGTGGCCATGAGTCTCATCAGAGAGTCCGGA 180
QY 181 GAGGCGCTGCGCTATACCGAGCGGATTTCAAGCGCGAGCGCGCAACTGCTGGCACC 240
DB 181 GAGGCGCTGCGCTATACCGAGCGGATTTCAAGCGCGAGCGCGCAACTGCTGGCACC 240
QY 241 GCCGACGAGTGTGGCGCGAGCGCTGATTTATGCTCAAGGTCAAAAGACGATAGCGGCG 300
DB 241 GCCGACGAGTGTGGCGCGAGCGCTGATTTATGCTCAAGGTCAAAAGACGATAGCGGCG 300

```

QY	301	GAATAGGCGCGCTCGACACAGGGGAGATCTTGTCACGT	CTTGCAATTGGCGCGCTCA	360
Db	301	GAATAGGCGCGCTCGACACAGGGGAGATCTTGTCACGT	CTTGCAATTGGCGCGCTCA	360
QY	361	CGTGCCTGCACCGATGCGTTGTTGGATTCCGGGCACACACGTC	CAATTGGCTACGACACGTC	420
Db	361	CGTGCCTGCACCGATGCGTTGTTGGATTCCGGGCACACACGTC	CAATTGGCTACGACACGTC	420
QY	421	CAGACCGCGGAGGGGCGACTACCCCTGCTTCCCGCATGAGACGAAAGTCCGCGTGCATC	480	
Db	421	CAGACCGCGGAGGGGCGACTACCCCTGCTTCCCGCATGAGACGAAAGTCCGCGTGCATC	480	
QY	481	GCCGCCAAGTTGGCGCTTACACCTGATGCGAACCCCAAGGGGGCCCGGTGCTCATG	540	
Db	481	GCCGCCAAGTTGGCGCTTACACCTGATGCGAACCCCAAGGGGGCCCGGTGCTCATG	540	
QY	541	GGCGGGGGTCCCGGGGTGGAACCCGGCCGACGTCGTGTGATTCGGCGCCGACCGCCGGC	600	
Db	541	GGCGGGGGTCCCGGGGTGGAACCCGGCCGACGTCGTGTGATTCGGCGCCGACCGCCGGC	600	
QY	601	TACAAACGACGCGCCGATCGCCAAAGGCGCATGCGGCGCGACCGCTTAGCGTTCTAGACATCAAC	660	
Db	601	TACAAACGACGCGCCGATCGCCAAAGGCGCATGCGGCGCGACCGCTTAGCGTTCTAGACATCAAC	660	
QY	661	ATCGACAAACTTCGGCAACTGACGCGCGAGTTCGTGCGGGCGGATCCAACTCGCTACTCA	720	
Db	661	ATCGACAAACTTCGGCAACTGACGCGCGAGTTCGTGCGGGCGGATCCAACTCGCTACTCA	720	
QY	721	TGCGGCTCAGAGACTGAGAGGTGCGCTCAAAAGTCCCGACCTGGGTGATTTGGGGCGGCTGTG	780	
Db	721	TGCGGCTCAGAGACTGAGAGGTGCGCTCAAAAGTCCCGACCTGGGTGATTTGGGGCGGCTGTG	780	
QY	781	GTGCGAGGGCGCCAGAGCCACCCAAATTAGCTTCGATTCACCTGTGCGGCATATGAAACCA	840	
Db	781	GTGCGAGGGCGCCAGAGCCACCCAAATTAGCTTCGATTCACCTGTGCGGCATATGAAACCA	840	
QY	841	GGTGGGATAGTGGGATATATGACATGACACAGGGCGGCGT	900	
Db	841	GGTGGGATAGTGGGATATATGACATGACACAGGGCGGCGT	900	
QY	901	ACGACCTCAGACCAACCCGACGTTGCGCGTGCACACACAGCGTGTGTTTACTCGTGGCGAAC	960	
Db	901	ACGACCTCAGACCAACCCGACGTTGCGCGTGCACACACAGCGTGTGTTTACTCGTGGCGAAC	960	
QY	961	ATGCGCGCGCTGGGTGTCGGAACAGCTGCACCTTAGCGGCTTGACCAACGCGAGATCCGCTAT	1020	
Db	961	ATGCGCGCGCTGGGTGTCGGAACAGCTGCACCTTAGCGGCTTGACCAACGCGAGATCCGCTAT	1020	
QY	1021	GTGCTCGAGCTTGGCGACATAGGCTGGCGGGGGCGCGTGGCGGTGCGAATCCGGGACTAAGCC	1080	
Db	1021	GTGCTCGAGCTTGGCGACATAGGCTGGCGGGGGCGCGTGGCGGTGCGAATCCGGGACTAAGCC	1080	
QY	1081	AAAGGCTCTTTCGACGACGACGAAGGGGGCTTACTGTCCGACAGGGGTGGCCACGACCTGGGG	1140	
Db	1081	AAAGGCTCTTTCGACGACGACGAAGGGGGCTTACTGTCCGACAGGGGTGGCCACGACCTGGGG	1140	
QY	1141	GTGCGGTCACACGAGCCCGGACGAGGTGCTGTGCGCTGATCTTCGGCGCTGTGTTAGCCGAG	1200	
Db	1141	GTGCGGTCACACGAGCCCGGACGAGGTGCTGTGCGCTGATCTTCGGCGCTGTGTTAGCCGAG	1200	
QY	1201	CACACGTGCGGAGTAAAGGAAGCATGATGTGCGCC	1236	
Db	1201	CACACGTGCGGAGTAAAGGAAGCATGATGTGCGCC	1236	
RESULT	13			
LOCUS	A89751	1235 bp	DNA	
DEFINITION	Sequence 8 from Patent W09832862.		PAT	22-JAN-2000
ACCESSION	A89751			
VERSION	A89751.1	GI:6738285		
KEYWORDS	unidentified.			

ORGANISM	unidentified
REFERENCE	1 (bases 1 to 1235)
AUTHORS	Flohe, L. and Singh, M.
TITLE	L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
JOURNAL	Patent: WO 9832862-A 30-JUL-1998;
FEATURES	FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
source	Location/Qualifiers 1..1235 /organism="unidentified" /db_xref="taxon:32644"
BASE COUNT	236 a 394 c 385 g 220 t
ORIGIN	
Query Match	98.7% Score 1220.8; DB 5; Length 1235;
Best Local Similarity	99.8% Pred. No. 5e-168;
Matches 1233; Conservative	0; Mismatches 2; Indels 1; Gaps 1;
1	ATCTTGGACATTAATTCGAACTTTCTTCTCACTGAGACGGTACAGTATCCGAGAGGGTAAATC 60
1	ATCTTGGACATTAATTCGAACTTTCTTCTCACTGAGACGGTACAGTATCCGAGAGGGTAAATC 60
61	ATGCGCGTGGTATTCGACGACGAGACCAAAACAAGAAATTCGGGTGGCCATCAACCCCG 120
61	ATGCGCGTGGTATTCGACGACGAGACCAAAACAAGAAATTCGGGTGGCCATCAACCCCG 120
121	GGCGGCGTCCGGCACTAACCCGCTGTGGCCATGAGGTGCTCATCCAGGCAAGTCCCGGA 180
121	GGCGGCGTCCGGCACTAACCCGCTGTGGCCATGAGGTGCTCATCCAGGCAAGTCCCGGA 180
181	GAGGCGTGGTATACGACGACGGATTTCAAGGCGGAGGCGGCACTGGTGGCGAC 240
181	GAGGCGTGGTATACGACGACGGATTTCAAGGCGGAGGCGGCACTGGTGGCGAC 240
241	GCCGACACAGTGTGGCGCCGACGCGTATTATTGCTCAAGGTCMAAGAACGATACGGCG 300
241	GCCGACACAGTGTGGCGCCGACGCGTATTATTGCTCAAGGTCMAAGAACGATACGGCG 300
301	GAATACGGGCGGCTGCGGACACGGGAGAGATCTGTTCACGTTCTTGCAATTTGGCGCGTCA 360
301	GAATACGGGCGGCTGCGGACACGGGAGAGATCTGTTCACGTTCTTGCAATTTGGCGCGTCA 360
361	CGTGTTCACCGATGCGTGTGTGGATTCGCGACACAGTCAATTTGGCTTACGACGCTC 420
361	CGTGTTCACCGATGCGTGTGTGGATTCGCGACACAGTCAATTTGGCTTACGACGCTC 420
421	CAGACCGCCGAGGCGCGCACTACCTCGTTGCCCGGATGAGCGAAGTGGCCGGTGCATC 480
421	CAGACCGCCGAGGCGCGCACTACCTCGTTGCCCGGATGAGCGAAGTGGCCGGTGCATC 480
481	GCGCGCCAGGTGGCGCTTACCACTGATGTGCAACCCCAAGGGGGCGGGGTGTGTATG 540
481	GCGCGCCAGGTGGCGCTTACCACTGATGTGCAACCCCAAGGGGGCGGGGTGTGTATG 540
541	GGCGGGGTGCCCGCGCTGAACCGGGCGCAGCTGTGTGATTCGGGCGCGACCGCGGC 600
541	GGCGGGGTGCCCGCGCTGAACCGGGCGCAGCTGTGTGATTCGGGCGCGACCGCGGC 600
601	TACACGAGCGCGCGATGGCCCAACCGCATGGGCGCGACCGCTTACGGTTCAGACATCAAC 660
601	TACACGAGCGCGCGATGGCCCAACCGCATGGGCGCGACCGCTTACGGTTCAGACATCAAC 660
661	ATCGACCAACTTCGGCACTCGACCGCCAGTGTTCGGCGCGGATTCACACATCGCTACTCA 720
661	ATCGACCAACTTCGGCACTCGACCGCCAGTGTTCGGCGCGGATTCACACATCGCTACTCA 720
721	TCGGCTTACGAGCTGAGAGGTTCGGTCAACAGTGCACCTGTGTGATTTGGGCGCTCTG 780
721	TCGGCTTACGAGCTGAGAGGTTCGGTCAACAGTGCACCTGTGTGATTTGGGCGCTCTG 780
781	GTGCGAGCGCCCAAGGACCCCAATTATCTCGAATTCACCTGTTCGGCATATCAACCA 840
781	GTGCGAGCGCCCAAGGACCCCAATTATCTCGAATTCACCTGTTCGGCATATCAACCA 840

```

Db 780 GTGCGAGGCGCCAGAGCAACCAATTAGTCTGCAATTCTTCCGCGCATATGAAACCA 839
QY 841 GGTGCGGACTAGTGTGATATATAGCATGACAGGCGGCGTGTTCGAAAGGCTCAGACCG 900
    |||
Db 840 GGTGCGGACTAGTGTGATATATAGCATGACAGGCGGCGTGTTCGAAAGGCTCAGACCG 899
QY 901 ACCACCTACGACCAACCGGAGCTTGGCGGTGACAGACAGCTGTTTACTGGTGGCGAAC 960
    |||
Db 900 ACCACCTACGACCAACCGGAGCTTGGCGGTGACAGACAGCTGTTTACTGGTGGCGAAC 959
QY 961 ATGCGCGCTGCTGCGGAGAGAGCTGACCTACGCGGTGACCAACGAGCATGCGGTAT 1020
    |||
Db 960 ATGCGCGCTGCTGCGGAGAGAGCTGACCTACGCGGTGACCAACGAGCATGCGGTAT 1019
QY 1021 GTGCTGAGCTTGGCGGACATGCTGGCGGCGGCGGTGCGGATCGGACATACCC 1080
    |||
Db 1020 GTGCTGAGCTTGGCGGACATGCTGGCGGCGGCGGTGCGGATCGGACATACCC 1079
QY 1081 AAAGGTCTTTCGAGCGACGAAAGGCGCTTACTGTCCGAAAGGCTGCGACGACCTGGGG 1140
    |||
Db 1080 AAAGGTCTTTCGAGCGACGAAAGGCGCTTACTGTCCGAAAGGCTGCGACGACCTGGGG 1139
QY 1141 GTGCGGCTTACCGAGCGCGCGGAGCGGTGCTGGCTGACTCTCGGCGCGCTGTTAGCGCGAG 1200
    |||
Db 1140 GTGCGGCTTACCGAGCGCGCGGAGCGGTGCTGGCTGACTCTCGGCGCGCTGTTAGCGCGAG 1199
QY 1201 CACAGTCGGGAGTAAGGGAAGCATGATGTGCGCC 1236
    |||
Db 1200 CACAGTCGGGAGTAAGGGAAGCATGATGTGCGCC 1235

RESULT 14
A87605 1245 bp DNA PAT 22-JAN-2000
LOCUS A87605
DEFINITION Sequence 2 from Patent WO9836089.
ACCESSION A87605
VERSION A87605.1 GI:6736245
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1245)
AUTHORS Flohe, L. and Singh, M.
TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL Patent: WO 9836089-A 20-AUG-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
source 1..1245
location/Qualifiers
/db_xref="taxon:32644"
/organism="unidentified"
BASE COUNT 238 a 398 c 387 g 222 t
ORIGIN

```

```

Query Match 98.6%; Score 1219.4; DB 5; Length 1245;
Best Local Similarity 99.4%; Pred. No. 8e-168;
Matches 1236; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

```

```

QY 235 GGCACCGCCGACAGGTGTGGCCGACGCTGATTATTTGCTCAAGGTCAAGAACGATA 294
    |||
Db 241 GGCACCGCCGACAGGTGTGGCCGACGCTGATTATTTGCTCAAGGTCAAGAACGATA 300
QY 295 GCGGCGGAATACGGCCCGCTGCGACACGGGAGATCTTTGTTACGTTCTTGCAATTGGCC 354
    |||
Db 301 GCGGCGGAATACGGCCCGCTGCGACACGGGAGATCTTTGTTACGTTCTTGCAATTGGCC 360
QY 355 GCGTACAGTGTGTCACCGATGCTTTTGTGTTCCGGGACCAAGTCAATTTGCTTAGAG 414
    |||
Db 361 GCGTACAGTGTGTCACCGATGCTTTGTTGTTCCGGGACCAAGTCAATTTGCTTAGAG 420
QY 415 ACCGTCCAGACCGCCGACAGGCGGACACTACCCCTGCTTCCCGGATGAGAGGAGTCCCGGT 474
    |||
Db 421 ACCGTCCAGACCGCCGACAGGCGGACACTACCCCTGCTTCCCGGATGAGAGGAGTCCCGGT 480
QY 475 CGACTCGGCGCCAGGTTGGCGCTTACCACCTGATGCGAACCAGAGGGGCGCGGTGTG 534
    |||
Db 481 CGACTCGGCGCCAGGTTGGCGCTTACCACCTGATGCGAACCAGAGGGGCGCGGTGTG 540
QY 535 CTGATGGGCGGGGTGGCGCGGCTGGAACCGGCGGAGCGTGGTGTGCGGCGCGGAGC 594
    |||
Db 541 CTGATGGGCGGGGTGGCGCGGCTGGAACCGGCGGAGCGTGGTGTGCGGCGCGGAGC 600
QY 595 GCGGCTACAAACGACGCGCGCATCGCAACGATGGGCGGACCGTTACGCTTAGAC 654
    |||
Db 601 GCGGCTACAAACGACGCGCGCATCGCAACGATGGGCGGACCGTTACGCTTAGAC 660
QY 655 ATCAACATCGACAACTTTCGGAACCTGACGCGGAGTCTTCCGCGCGGATCCACACTGC 714
    |||
Db 661 ATCAACATCGACAACTTTCGGAACCTGACGCGGAGTCTTCCGCGCGGATCCACACTGC 720
QY 715 TACTCATGGGCTTACGAGCTGAGAGGTGGCGTCAAGGTGGCGGACCTGATTTGGGGCC 774
    |||
Db 721 TACTCATGGGCTTACGAGCTGAGAGGTGGCGTCAAGGTGGCGGACCTGATTTGGGGCC 780
QY 775 GTCTGTGTGCCAGGCGCGCAAGGACCCCAATTACTCTGAAATTAATCTTGGCGCATATG 834
    |||
Db 781 GTCTGTGTGCCAGGCGCGCAAGGACCCCAATTACTCTGAAATTAATCTTGGCGCATATG 840
QY 835 AAACAGGTGGGCTACTGCGATATAGCCATGACAGGCGGCTGTTGGAAGGCTCA 894
    |||
Db 841 AAACAGGTGGGCTACTGCGATATAGCCATGACAGGCGGCTGTTGGAAGGCTCA 900
QY 895 CGACCGAGACCTACGACACCCGACGCTTGGCGGTGACAGACAGCTGTTTACTGGCTG 954
    |||
Db 901 CGACCGAGACCTACGACACCCGACGCTTGGCGGTGACAGACAGCTGTTTACTGGCTG 960
QY 955 GCGAAGCATGCCGCGCTCGCTGCTGCAAGACGTGACCTACGCGGTGACCAACGCGAGATG 1014
    |||
Db 961 GCGAAGCATGCCGCGCTCGCTGCTGCAAGACGTGACCTACGCGGTGACCAACGCGAGATG 1020
QY 1015 CCGTATGTGCTGAGCTTGGCGGACATGCTGCGGCGGCGGCGTCCCGGTGCAATCCGCA 1074
    |||
Db 1021 CCGTATGTGCTGAGCTTGGCGGACATGCTGCGGCGGCGGCGTCCCGGTGCAATCCGCA 1080
QY 1075 GTAGCAAAAGGCTTTCGAGCGACGAAAGGCGCTTACTGTTGGAAGGCGGAGGAG 1134
    |||
Db 1081 GTAGCAAAAGGCTTTCGAGCGACGAAAGGCGCTTACTGTTGGAAGGCGGAGGAG 1140
QY 1135 CTGGGGGTGGCGCTTACCGAGACCGCGGACGCTGCTGGCTACTCTGCGCGCTGTTAC 1194
    |||
Db 1141 CTGGGGGTGGCGCTTACCGAGACCGCGGACGCTGCTGGCTACTCTGCGCGCTGTTAC 1200
QY 1195 GCGGAGCACAGCTGCGGAGTAAGGGAAGCATGATGTGCGCGG 1237
    |||
Db 1201 GCGGAGCACAGCTGCGGAGTAAGGGAAGCATGATGTGCGCGG 1243

RESULT 15
A89745 1245 bp DNA PAT 22-JAN-2000
LOCUS A89745

```

DEFINITION Sequence 2 from Patent WO9832862.
 ACCESSION A89745
 VERSION A89745.1 GI:6738279
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1245)
 FLOHE, L. and Singh, M.
 L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
 JOURNAL Patent: WO 9832862-A 30-JUL-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 Location/Qualifiers
 1..1245
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 238 a 398 c 387 g 222 t
 ORIGIN

Query Match 98.6%; Score 1219.4; DB 5; Length 1245;
 Best Local Similarity 99.4%; Pred. No. 8e-168;
 Matches 1236; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 1 ATCTGAGATTAAATCGAACTTTCTTCACTGAGAGCTAGATGAGAGGGTATC 60
 Db 1 ATCTGAGATTAAATCGAACTTTCTTCACTGAGAGCTAGATGAGAGGGTATC 60
 QY 61 ATGGCGGTGATTCGACGAGACCAAAACAAG-ATTCGGGGGCGCATC 114
 Db 61 ATGGCGGTGATTCGACGAGACCAAAACAAG-ATTCGGGGGCGCATC 120
 QY 115 ACCCGGCGGCGTGGGAGAACTAACCGGCGCATGAGGTCATCATCAGGAGCT 174
 Db 121 ACCCGGCGGCGTGGGAGAACTAACCGGCGCATGAGGTCATCATCAGGAGCT 180
 QY 175 GCGGAGAGGGGCTGGCTATACCGAGCGGATTTCAAGGCGGCGCGCAATGCTC 234
 Db 181 GCGGAGAGGGGCTGGCTATACCGAGCGGATTTCAAGGCGGCGCGCAATGCTC 240
 QY 235 GGCACCGCGGCGAGGTGCGGCGGCGCTGATTTATGCTCAAGGTCAAAGAACGATA 294
 Db 241 GGCACCGCGGCGAGGTGCGGCGGCGCTGATTTATGCTCAAGGTCAAAGAACGATA 300
 QY 295 GCGGCGGATACGCGCGGCGCTGCGACAGCGGAGATCTGTCAGCTTCTTCATTTGGCC 354
 Db 301 GCGGCGGATACGCGCGGCGCTGCGACAGCGGAGATCTGTCAGCTTCTTCATTTGGCC 360
 QY 355 GCGTCAGCTGCTTGCACCGATGCGTGTGATTCGCGACCGATTCATTCCTACGAG 414
 Db 361 GCGTCAGCTGCTTGCACCGATGCGTGTGATTCGCGACCGATTCATTCCTACGAG 420
 QY 415 ACCGTCCGACCGCGGAGGCGCACTACCGCTGCTGCGCGGATGAGGAAAGTGGCGGT 474
 Db 421 ACCGTCCGACCGCGGAGGCGCACTACCGCTGCTGCGCGGATGAGGAAAGTGGCGGT 480
 QY 475 CGACTCGCGCGCGAGGTTGGGCTTACCACTGATGCGAACCAAGGGGGCGCGGTGTG 534
 Db 481 CGACTCGCGCGCGAGGTTGGGCTTACCACTGATGCGAACCAAGGGGGCGCGGTGTG 540
 QY 535 CTGATGGGCGGCGGCGCGGCGGCGGATGCGAACCGCGCACTGCTGATGCGCGCGGCAAC 594
 Db 541 CTGATGGGCGGCGGCGGCGGCGGCGGATGCGAACCGCGCACTGCTGATGCGCGCGGCAAC 600
 QY 595 GCGGCGTACAAAGCGAGCGCGATGCGCAAGCGATGCGGCGGCGGCGGATGCGGTTCTAGAC 654
 Db 601 GCGGCGTACAAAGCGAGCGCGATGCGCAAGCGATGCGGCGGCGGCGGATGCGGTTCTAGAC 660
 QY 655 ATCAACATCGAACAATTCGGAACCTGAGCGCGGAGTTCTGCGCGGCGGATGCGGTTCTAGAC 714
 Db 661 ATCAACATCGAACAATTCGGAACCTGAGCGCGGAGTTCTGCGCGGCGGATGCGGTTCTAGAC 720
 QY 715 TACTCATGCGGCTGAGAGCTGAGGTTGCGGTCAAAGCTGCGGACCTGGTATGGGGCC 774

Db 721 TACTCATGCGGCTGAGAGCTGAGGTTGCGGTCAAAGCTGCGGACCTGGTATGGGGCC 780
 QY 775 GTCTGTGCTCCAGGCGCGCAAGCGCAACCAATTAAGTCTGCAATTCAGTTGCGCGCATATG 834
 Db 781 GTCTGTGCTCCAGGCGCGCAAGCGCAACCAATTAAGTCTGCAATTCAGTTGCGCGCATATG 840
 QY 835 AAACGAGTGGGTAGTGGGTGATATAGCATGACAGGCGGCGGCTTTGCAAGGCTCA 894
 Db 841 AAACGAGTGGGTAGTGGGTGATATAGCATGACAGGCGGCGGCTTTGCAAGGCTCA 900
 QY 895 CGACCGACACCTAGACCAACCGGAGCTTGGCGGTGCGACGACGCTGTTTACTGCGTG 954
 Db 901 CGACCGACACCTAGACCAACCGGAGCTTGGCGGTGCGACGACGCTGTTTACTGCGTG 960
 QY 955 GCGAACAATGCGGCGGCGGCTGCGGAGAGAGTGGAGTACGCGCTGACCAAGCGAGATG 1014
 Db 961 GCGAACAATGCGGCGGCGGCTGCGGAGAGAGTGGAGTACGCGCTGACCAAGCGAGATG 1020
 QY 1015 CCGTATGTGCTCGAGCTTGCGGACATGCTGCGGCGGCGGCGGCGGCGGATGCGGCA 1074
 Db 1021 CCGTATGTGCTCGAGCTTGCGGACATGCTGCGGCGGCGGCGGCGGCGGATGCGGCA 1080
 QY 1075 CTAGCCAAAGCTTTTCGACGACGCAAGGCGGTTACTGTCCGAAGGCTGCGCACCGAC 1134
 Db 1081 CTAGCCAAAGCTTTTCGACGACGCAAGGCGGTTACTGTCCGAAGGCTGCGCACCGAC 1140
 QY 1135 CTGCGGGGTCGCTTACCGAGTCCGCGACGCTGCTGAGCTCGCGGCGCTGCTGTTAC 1194
 Db 1141 CTGCGGGGTCGCTTACCGAGTCCGCGACGCTGCTGAGCTCGCGGCGCTGCTGTTAC 1200
 QY 1195 GCGGACGACGCTGCGGAGTAAAGGAAAGCATGATGCTGGCGG 1237
 Db 1201 GCGGACGACGCTGCGGAGTAAAGGAAAGCATGATGCTGGCGG 1243

Search completed: June 22, 2000, 14:52:44
 Job time: 17579 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 15:06:27 ; Search time 458.59 seconds
(without alignments)
674.868 Million cell updates/sec

Title: US-09-362-485-4

Perfect score: 1237
Sequence: 1 ATCTTGCAGATTAAATCGAAC.....GGAGCGATGATGCGCCG 1237

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameter: 6.3170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1219.4	98.6	1245	1 V49626	Mycobacterium tube
2	1219.4	98.6	1260	1 V49510	Mycobacterium sp.
3	1219.4	98.6	1260	1 V49625	Mycobacterium tube
4	467.6	37.8	682	1 V49511	Mycobacterium mari
5	277.2	22.4	1125	1 N91423	Sequence of heat-r
6	212.4	17.2	28171	1 V52155	Streptococcus pneu
7	178.4	14.4	9280	1 V74442	Streptococcus pneu
8	94.8	7.7	1074	1 T67971	H. pylori membrane
9	69.2	5.6	544	1 V43039	Streptococcus pneu
10	57.6	4.7	31096	1 V74370	Streptococcus pneu
11	50.8	4.1	390	1 Q21833	Randomising oligon
12	50.8	4.1	390	1 Q36859	PCR primer for 5'
13	50.4	4.1	1413	1 X34205	Mycobacterium spec
14	50.4	4.1	1722	1 X34206	Mycobacterium spec
15	50.4	4.1	15872	1 T68715	Streptomycetes venez
16	49	4.0	1269	1 Q03065	Sequence encoding
17	47.6	3.8	3346	1 T93610	Mycobacterium tube
18	47.6	3.8	28598	1 T06769	Sorangium cellulos
19	47.6	3.8	28958	1 T89956	Sorangium cellulos
20	47.6	3.8	49377	1 V05287	The soraphen biosy
21	46.4	3.8	985	1 V44439	Mycobacterium tube
22	46.4	3.8	985	1 V64548	M. tuberculosis im
23	46	3.7	1620	1 Q22482	griDL-1 gene codin
24	46	3.7	2582	1 T73117	Actinoplanes sp. a
25	46	3.7	2668	1 Q22485	griDL-1 gene. Reco
26	46	3.7	17955	1 V56642	Actinoplanes sp. a
27	45.4	3.7	5392	1 Q64203	snb gene encoding
28	45.4	3.7	5392	1 Q64201	Sequence comprisin
29	42.4	3.4	1833	1 Q64206	snb gene encoding
30	42.4	3.4	2185	1 V84066	Clone p5-4 encodin
31	42.4	3.4	2186	1 Q09365	S. lividans protea
32	42.4	3.4	24379	1 T93095	Streptomycetes roseo
33	42.4	3.4	24379	1 V25925	Streptomycetes roseo
34	42.2	3.4	329	1 V44425	Mycobacterium tube

35	42.2	3.4	329	1 V64534	M. tuberculosis im
36	42	3.4	3946	1 T93610	Mycobacterium tube
37	42	3.4	4018	1 Q63879	Polyhydroxyalkanoa
38	41.8	3.4	882	1 V44403	Mycobacterium tube
39	41.8	3.4	882	1 V64512	M. tuberculosis im
40	41.6	3.4	12036	1 Q04668	FHA structural gen
41	41.2	3.3	2414	1 Q05926	Sequence encoding
42	41	3.3	29879	1 Q46806	eryA region of S.
43	40.8	3.3	2151	1 Q76252	HSV-2 protease, IC
44	40.8	3.3	2151	1 Q76261	HSV-2 protease/ICP
45	40.8	3.3	2472	1 Q84671	HSV-2 UL26 gene. N

ALIGNMENTS

RESULT	ID	Score	Query Match	Length	DB ID	Description
1	V49626	98.6%	99.4%	1245	1 V49626	Mycobacterium tube
2	V49626	98.6%	99.4%	1260	1 V49510	Mycobacterium sp.
3	V49626	98.6%	99.4%	1260	1 V49625	Mycobacterium tube
4	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
5	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
6	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
7	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
8	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
9	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
10	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
11	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
12	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
13	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
14	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
15	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
16	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
17	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
18	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
19	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
20	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
21	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
22	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
23	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
24	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
25	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
26	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
27	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
28	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
29	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
30	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
31	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
32	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
33	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
34	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
35	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
36	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
37	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
38	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
39	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
40	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
41	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
42	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
43	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
44	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari
45	V49626	98.6%	99.4%	1260	1 V49511	Mycobacterium mari

235 GGCACCGCCGACAGGTTGGGCGGCGGATTTATGCTCAAGGTCAAAGAACGATA 294
 241 GGCACCGCCGACAGGTTGGGCGGCGGATTTATGCTCAAGGTCAAAGAACGATA 300
 295 GCGGCGGAATACGGCGCGCTGCACACGGGCGAGATCTTTCAGCTTTCGATTTGGCC 354
 301 GCGGCGGAATACGGCGCGCTGCACACGGGCGAGATCTTTCAGCTTTCGATTTGGCC 360
 355 GCGTCAAGGCTTGCACGCAATGCTGTGGATTCGGGACCGGCAATTCGCTCGAG 414
 361 GCGTCAAGGCTTGCACGCAATGCTGTGGATTCGGGACCGGCAATTCGCTCGAG 420
 415 ACCGTCAGACCGCCGACGCGGCGGATACCGCTGCTTCCCGGATGAGCGAAGTCGGGT 474
 421 ACCGTCAGACCGCCGACGCGGCGGATACCGCTGCTTCCCGGATGAGCGAAGTCGGGT 480
 475 CGACTCGCGCGGCGGATGCTTACGACGCTGATCGGAAACCGGCGGCGGCTG 534
 481 CGACTCGCGCGGCGGATGCTTACGACGCTGATCGGAAACCGGCGGCGGCTG 540
 535 CTGATGGGCGGCGGCTGCGGCGGCTGCAACCGGCGGCGGATGCTGCGGCGGCGG 594
 541 CTGATGGGCGGCGGCTGCGGCGGCTGCAACCGGCGGCGGATGCTGCGGCGGCGG 600
 595 GCGGCGTACACGACCGCGGATGCGGCGGCGGATGCGGCGGCGGATGCGGCGGCGG 654
 601 GCGGCGTACACGACCGCGGATGCGGCGGCGGATGCGGCGGCGGATGCGGCGGCGG 660
 655 ATCAATCATGACAAATCTTGGGCAATCTGAGCGGCGGATGCTGCGGCGGCGGATG 714
 661 ATCAATCATGACAAATCTTGGGCAATCTGAGCGGCGGATGCTGCGGCGGCGGATG 720
 715 TACTCATCGGCGCTACGAGCTCGAGGTCGCTCAAAAGTCGCGGCGGCTGATGGGCGC 774
 721 TACTCATCGGCGCTACGAGCTCGAGGTCGCTCAAAAGTCGCGGCGGCTGATGGGCGC 780
 775 GTCCTGTCGAGGCGGCGCAAGGCGCAATTAATGTCGAATCAATTCGCGGCGGATG 834
 781 GTCCTGTCGAGGCGGCGCAAGGCGCAATTAATGTCGAATCAATTCGCGGCGGATG 840
 835 AAACGAGTGGGCTGCTGATGATATGATGATGATGATGATGATGATGATGATGATG 894
 841 AAACGAGTGGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 900
 895 CGACCGACCACTACGACCAACCGGAGCTTGGCGGCTGCAAGACGCTGTTTACTGCTG 954
 901 CGACCGACCACTACGACCAACCGGAGCTTGGCGGCTGCAAGACGCTGTTTACTGCTG 960
 955 GCGAATCATGCGGCTGCTGATGCTGCGGAGAGGCTGACCTGACGCGTGAACGAGGATG 1014
 961 GCGAATCATGCGGCTGCTGATGCTGCGGAGAGGCTGACCTGACGCGTGAACGAGGATG 1020
 1015 CCGTATGCTGCTGAGCTTGGCGGACATGCTGCGGCGGCGGCTGCGGCGGATG 1074
 1021 CCGTATGCTGCTGAGCTTGGCGGACATGCTGCGGCGGCGGCTGCGGCGGATG 1080
 1075 CTAGCCAAAGGCTTTCGACGACGAAAGGCGGCTTCTGTCGAAAGGCGGCGGCGG 1134
 1081 CTAGCCAAAGGCTTTCGACGACGAAAGGCGGCTTCTGTCGAAAGGCGGCGGCGG 1140
 1135 CTGGGGGCTGCTGACGACGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1194
 1141 CTGGGGGCTGCTGACGACGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 1195 GCGGACACACGTCGGGATTAAGGGAAGGATGATGCTGCGGCGG 1237
 1201 GCGGACACACGTCGGGATTAAGGGAAGGATGATGCTGCGGCGG 1243

RESULT 2
 ID V49510 standard: DNA, 1260 BP.
 AC V49510;

DT 20-OCT-1998 (first entry)
 DE Mycobacterium sp. Aladh DNA.
 KM Alanine dehydrogenase; Aladh; ADH; diagnosis; tuberculosis; pathogen;
 KW swimmers' disease; vaccine; epidemic; infection; identification; ss.
 OS Mycobacterium sp.
 PN M09832862-A2.
 PD 30-JUL-1998.
 PF 29-JAN-1998; E00484.
 PR 29-JAN-1997; EP-101339.
 PA (FLOH/) FLOHE L.
 PI Flohe L, Butler B, Kolk A, Singh M;
 DR WPI: 98-427958/36.
 PR Nucleic acid encoding alanine dehydrogenase of Mycobacterium marinum
 PT - used for diagnosis of tuberculosis and other mycobacterial
 PT diseases, also for treatment and prevention, for drug screening and
 PT for bio-transformation
 PS Disclosure: Page 11: 57pp. German.
 CC This sequence encodes an alanine dehydrogenase (ADH) protein isolated
 CC from a Mycobacterium sp. This protein is used to diagnose tuberculosis
 CC and other mycobacterial infections (including 'swimmers' disease', caused
 CC by M. marinum, a fish pathogen) in humans or animals. The protein can
 CC also be used for control of epidemics and for vaccination, to screen for
 CC agents with anti-mycobacterial activity, and in bio-transformations that
 CC are specific for L-alanine. Also mycobacteria can be identified by
 CC analysis of genomic ADH sequences. ADH is an antigen that is secreted
 CC early during infection.
 SQ Sequence 1260 BP; 243 A; 403 C; 389 G; 225 T;

Query Match 98.6%; Score 1219.4; DB 1; Length 1260;
 Best Local Similarity 99.4%; Pred. No. 1.4e-245;
 Matches 1236; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 1 ATCTTGACGATTAATCGAATCTTCTGCTACTGAGCGTACGATGAGGCGGTATC 60
 DB 16 ATCTTGACGATTAATCGAATCTTCTGCTACTGAGCGTACGATGAGGCGGTATC 75
 QY 61 ATGCGCGTGGTATTCGACCGGAGCAAAAACAACG-----AATTCGGGTGGCCATC 114
 DB 76 ATGCGCGTGGTATTCGACCGGAGCAAAAACAACGATTCGAATTCGGGTGGCCATC 135
 QY 115 ACCCGCGCGGCGGCGGCGGAGTAAACCGGTCGCGGATGAGGTCGATCCAGGCGGT 174
 DB 136 ACCCGCGCGGCGGCGGCGGAGTAAACCGGTCGCGGATGAGGTCGATCCAGGCGGT 195
 QY 175 GCGGAGAGGCGCTCGGCTATCACCGGCGGATTTCAAGGCGGCGGCGGCACTGCTC 234
 DB 196 GCGGAGAGGCGCTCGGCTATCACCGGCGGATTTCAAGGCGGCGGCGGCACTGCTC 255
 QY 235 GGCACCGCGGCGGCGGCGGAGTGGGCTGAGCGCTATTTATGCTCAAGGTCGAAGAACGATA 294
 DB 256 GGCACCGCGGCGGCGGCGGAGTGGGCTGAGCGCTATTTATGCTCAAGGTCGAAGAACGATA 315
 QY 295 GCGGCGGAATACGGCGCGCTGCGACGCGGCGGATCTTGTGCTCAAGGTCGAAGAACGATA 354
 DB 316 GCGGCGGAATACGGCGCGCTGCGACGCGGCGGATCTTGTGCTCAAGGTCGAAGAACGATA 375
 QY 355 GCGTACGCTGCTTGCACGCGGATGCTGTTGGATTCCGCGACACGATGATTTGCTACGAG 414
 DB 376 GCGTACGCTGCTTGCACGCGGATGCTGTTGGATTCCGCGACACGATGATTTGCTACGAG 435
 QY 415 ACCGTCAGACCGCCGACGCGGCGGCGGATACCGCTGCTTCCCGGATGAGCGAAGTCGGGT 474
 DB 436 ACCGTCAGACCGCCGACGCGGCGGCGGATACCGCTGCTTCCCGGATGAGCGAAGTCGGGT 495
 QY 475 CGACTCGCGCGGCGGCGGCTTACCACTGATGGAACCCAAAGGCGGCGGCGGTG 534
 DB 496 CGACTCGCGCGGCGGCGGCTTACCACTGATGGAACCCAAAGGCGGCGGCGGTG 555
 QY 535 CTGATGGGCGGCGGCGGCTGCAACCGGCGGCGGATGCTGATGCGGCGGCGGCGG 594
 DB 556 CTGATGGGCGGCGGCGGCTGCAACCGGCGGCGGATGCTGATGCGGCGGCGGCGG 615

QY 595 GCCGCTACAAAGAGCCCGCATCGCCAAAGCGATGGGCGGACGCTTACGGTTCTAGAC 654
 DB 616 GCCGCTACAAAGAGCCCGCATCGCCAAAGCGATGGGCGGACGCTTACGGTTCTAGAC 675
 QY 655 ATCAACATCGCAAACTTGGCACTGCAAGCCGAGTTCTCGGGCGGATCCACACTGCC 714
 DB 676 ATCAACATCGCAAACTTGGCACTGCAAGCCGAGTTCTCGGGCGGATCCACACTGCC 735
 QY 715 TACTCATCGGCTACGAGCTGAGAGGTGCGTCAAAAGCTGCGGACCTGATGGGGGCC 774
 DB 736 TACTCATCGGCTACGAGCTGAGAGGTGCGTCAAAAGCTGCGGACCTGATGGGGGCC 795
 QY 775 GTCTGTGTCAGAGGCGCCAAAGCAACCAATTAGTCTCGAAATTCATTGCGCGCATATG 834
 DB 796 GTCTGTGTCAGAGGCGCCAAAGCAACCAATTAGTCTCGAAATTCATTGCGCGCATATG 855
 QY 835 AAACAGGTGGGTACTGATGATATACCATCGACCGGCGGCTGTTTGAAGGCTCA 894
 DB 856 AAACAGGTGGGTACTGATGATATACCATCGACCGGCGGCTGTTTGAAGGCTCA 915
 QY 895 CGACCGACCTACGAGCAACCGACGTTCCGCGGTGACACAGCTGTTTACTGCGGTG 954
 DB 916 CGACCGACCTACGAGCAACCGACGTTCCGCGGTGACACAGCTGTTTACTGCGGTG 975
 QY 955 GCGAATATGCGGCTGCGGTGCGGAAAGCTGACCTACGCGCTGACCAACGCAAGATG 1014
 DB 976 GCGAATATGCGGCTGCGGTGCGGAAAGCTGACCTACGCGCTGACCAACGCAAGATG 1035
 QY 1015 CCGATGTCGTCGAGCTTGGCGACCATGGTGGGCGGCGGTGCGGCTGCAATCGGCA 1074
 DB 1036 CCGATGTCGTCGAGCTTGGCGACCATGGTGGGCGGCGGTGCGGCTGCAATCGGCA 1095
 QY 1075 CTAGCCAAAGGTCTTTGAGCGACGCAAGGCGGCTTACTGTCGAAGCGGTGGCCACGAC 1134
 DB 1096 CTAGCCAAAGGTCTTTGAGCGACGCAAGGCGGCTTACTGTCGAAGCGGTGGCCACGAC 1155
 QY 1135 CTGGGGGGCGGTCACGAGCCCGGCGGCGGCTGTCGCTGACCTGCGGCGCTGTTAC 1194
 DB 1156 CTGGGGGGCGGTCACGAGCCCGGCGGCGGCTGTCGCTGACCTGCGGCGCTGTTAC 1215
 QY 1195 GCCGAGCAGACGTCGAGAGTAAAGGAAAGCGATGATGTCGGCG 1237
 DB 1216 GCCGAGCAGACGTCGAGAGTAAAGGAAAGCGATGATGTCGGCG 1258

RESULT 3
 V49625
 ID V49625 standard; DNA; 1260 BP.
 AC V49625;
 DT 20-NOV-1998 (first entry)
 DE Mycobacterium tuberculosis Alamine dehydrogenase.
 KW ss: Alamine dehydrogenase; tuberculosis; TB; mycobacteria; L-alanine.
 OS Mycobacterium tuberculosis.
 PN MO9836089-A2.
 PD 20-AUG-1998.
 PE 29-JAN-1998; E00483.
 PR 29-JAN-1997; EP-101338.
 PA (FLOH/) FLOHE L.
 PI FLOHE L, Hutter B, Kolk A, Singh M;
 PR WPL; 98-457123/39.
 PT Diagnosis of, e.g. tuberculosis from alamine dehydrogenase activity
 - useful for, e.g. for diagnosis, differentiation of strains,
 monitoring vaccination and identification of mycobacterial
 inhibitors
 PS Claim 13; Fig 2.3; 55pp; German.
 CC The Mycobacterium tuberculosis alamine dehydrogenase (AladH) is in
 the production of kits for diagnosing tuberculosis (TB) and other
 CC mycobacterial infections in humans or animals. Kits are used for direct
 CC diagnosis of TB on clinical samples (e.g. body fluids) and can
 CC differentiate between pathogenic and non-virulent strains, e.g. for
 CC identifying pathogens of the M. tuberculosis (M. t.) complex. The kit may
 CC also be used to identify substances that inhibit mycobacteria, for
 CC combatting epidemics and for vaccination follow-up. Oligonucleotides

CC derived from AladH are used similarly in diagnostic hybridisation tests,
 CC also for culture confirmation of isolated strains and for chromosome
 CC fingerprinting to detect/differentiate between mycobacteria, and for
 CC L-alanine-specific biotransformation reactions. AladH is an early
 CC antigen, present extracellularly after only a few days of growth, making
 CC it an ideal drug target.
 SQ Sequence 1260 BP; 243 A; 403 C; 389 G; 225 T;

Query Match 98.5%; Score 1219.4; DB 1; Length 1260;
 Best Local Similarity 99.4%; Pred. No. 1.4e-245;
 Matches 1236; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 1 ATCTGAGATTATGCACTTCTTCATCTGAGCGTACCTATCGAAGGGGTAATC 60
 DB 16 ATCTGAGATTATGCACTTCTTCATCTGAGCGTACCTATCGAAGGGGTAATC 75
 QY 61 ATGCGGCTCGGTATTCGACCGAGACCAAAACAGG-----AATTCGGGTGGCATC 114
 DB 76 ATGCGGCTCGGTATTCGACCGAGACCAAAACAGGATTCCATTCCGGGTGGCATC 135
 QY 115 ACCCGGCGGCGTGGGGAATTAACCGCTGCTGCGCATAGAGTCTATCGAGGAGT 174
 DB 136 ACCCGGCGGCGTGGGGAATTAACCGCTGCTGCGCATAGAGTCTATCGAGGAGT 195
 QY 175 GCCGAGAGGGCTCGGCTATCACCGGCGGATTCAGAGGCGGCGGCGGCAACTGGTC 234
 DB 196 GCCGAGAGGGCTCGGCTATCACCGGCGGATTCAGAGGCGGCGGCGGCAACTGGTC 255
 QY 235 GGCACGCGCAGCAGGTGTTGGGCGGAGCTGATTTATGCTCAAGGTCAAGAACGATA 294
 DB 256 GGCACGCGCAGCAGGTGTTGGGCGGAGCTGATTTATGCTCAAGGTCAAGAACGATA 315
 QY 295 GCGGCGGAATAGCGCGCGCGCGGCGGCGGAGATCTTGTTCAGCTTCTTGGCGC 354
 DB 316 GCGGCGGAATAGCGCGCGCGCGGCGGAGATCTTGTTCAGCTTCTTGGCGC 375
 QY 355 GCGTCAAGTCTTTCAGCCGATGCTGTTGATTCGCGGCGGCGGCGGCGGCGGAG 414
 DB 376 GCGTCAAGTCTTTCAGCCGATGCTGTTGATTCGCGGCGGCGGCGGCGGCGGAG 435
 QY 415 ACCGTCAGACCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 474
 DB 436 ACCGTCAGACCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 495
 QY 475 CGACTCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 534
 DB 496 CGACTCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 555
 QY 535 CTGATGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 594
 DB 556 CTGATGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 615
 QY 595 GCGGCGTACAGCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 654
 DB 616 GCGGCGTACAGCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 675
 QY 655 ATCAACATCGCAAACTTGGCACTGCAAGCCGAGTTCTCGAAATTCATTGCGCGCAT 714
 DB 676 ATCAACATCGCAAACTTGGCACTGCAAGCCGAGTTCTCGAAATTCATTGCGCGCAT 735
 QY 715 TACTCATCGGCTACGAGCTGAGAGGTGCGTCAAAAGCTGCGGACCTGATGGGGCC 774
 DB 736 TACTCATCGGCTACGAGCTGAGAGGTGCGTCAAAAGCTGCGGACCTGATGGGGCC 795
 QY 775 GTCTGTGTCAGAGGCGCCAAAGCAACCAATTAGTCTCGAAATTCATTGCGCGCAT 834
 DB 796 GTCTGTGTCAGAGGCGCCAAAGCAACCAATTAGTCTCGAAATTCATTGCGCGCAT 855
 QY 835 AAACAGGTGGGTACTGATGATATACCATCGACCGGCGGCTGTTTGAAGGCTCA 894
 DB 856 AAACAGGTGGGTACTGATGATATACCATCGACCGGCGGCTGTTTGAAGGCTCA 915

Matches 604; Conservative 0; Mismatches 458; Indels 12; Gaps 4;

```

QY 61 ATGCGCGTGGGATTCGCGACCGGACCAAAACAAAGATTCGGGTGGCCATACCCCG 120
DB 1 ATGAAAGATGGGATTCGCGACCAAAACAAAGATTCGGGTGGCCATACCCCG 60
QY 121 GCGGCGTGGCGAAATCAACCCGTCGTCATGAGTCTCT--CATCCAGGAGGTGCG 177
DB 61 GCGGCGTGGCGAAATTCGCGACCGGACCAAAAGCGGATGAGTCTCTATGAGAGCAAGCG 120
QY 178 GAGAGGCGTGGCGATTCACCGAGCGGATTCAGCGCGGACCGCGCAACTGTGTGCG 237
DB 121 GCGGCGTGGCGGATTCGCGACCGGATTCAGAGTAAAGAACCGCGGACGATGATCTGTC 180
QY 238 ACCGCGGACGAGTGTGGCGCGGATTCGTCGATTCGTCGATTCGTCGATTCGTCGATTC 297
DB 181 CGAAGTGGAGAGATGCTTGGAGCGGCGGATGCTTGAAGTGAAGCGCGGCTGCT 240
QY 298 GCGGATTCGCGCGCGCGCGACACGCGGAGATTCGTCGATTCGTCGATTCGTCGATTCG 357
DB 241 CGAGATTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 358 TCACGCTGTCGACCGGATGCTTGGATTCGCGGACCAACGTCATTCGTCGACGAC 417
DB 301 GCGGACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY 418 GTCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 477
DB 361 GTCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
QY 478 CTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 537
DB 421 ATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
QY 538 ATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 597
DB 481 CTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
QY 598 GCGTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 657
DB 541 GCGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 658 AACATGACAACTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 717
DB 601 AACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
QY 718 TCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 777
DB 661 TCACACTGCGATTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
QY 778 CTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 837
DB 721 TGTATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 777
QY 838 CAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 894
DB 778 CCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 837
QY 895 CGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 954
DB 838 CGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 897
QY 955 GCGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1014
DB 898 GCGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 954
QY 1015 CCGTATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1074
DB 955 CCGTATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1014
QY 1075 CAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1128
DB 1015 CAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1068

```

RESULT 6

V52155/c
ID V52155 standard; DNA: 28171 BP.
AC V52155;
DR 23-OCT-1998 (first entry)
DE Streptococcus pneumoniae genome fragment SEQ ID NO:22.
KW Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
OS computer readable medium; vaccine; pharmaceutical composition; ds.
PN MO9818931-A2.
PD 07-MAY-1998.
PF 30-OCT-1997: U19588.
PR 31-OCT-1996: US-029960.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M,
PI Kunsch CA, Rosen CA;
DR WPI: 98-272225/24.
PT Computer-readable medium with recorded Streptococcus pneumoniae
PT polynucleotide sequences - useful in diagnostic kits and assays, and
PT pharmaceutical compositions and vaccines for Streptococcus
PT pneumoniae
PS Claim 1: Page 273-289; 1409pp; English.
CC The present invention describes a computer readable medium which has
CC the nucleotide sequences SEQ ID NO:1 to 391 (V52134 to V52524) recorded
CC on it, or a representative fragment or a sequence at least 95% identical
CC to SEQ ID NO:1 to 391. The nucleotide sequences depicted in SEQ ID NO:1
CC to 391 (V52134 to V52524) are genomic fragments from Streptococcus
CC pneumoniae. The present invention also describes an isolated nucleic acid
CC molecule encoding a homologue of any of the fragments of the S. pneumoniae
CC genome (SEQ ID NO:1 to 391) where the nucleic acid molecule is produced
CC by a process comprising: (a) screening a genomic DNA library using as a
CC probe a target sequence defined by any of the sequences in SEQ ID NO:1
CC to 391, identifying members of the library which contain sequences
CC that hybridize to the target sequence and isolating the nucleic acid
CC molecules from the members; or (b) isolating mRNA, DNA or cDNA produced
CC from an organism, amplifying nucleic acid molecules whose nucleotide
CC sequence is homologous to amplification primers derived from the
CC fragment of the S. pneumoniae genome to prime the amplification and
CC isolating the amplified sequences. The computer readable medium can be
CC used in a computer-based system for identifying fragments of the
CC S. pneumoniae genome of commercial importance, or expression modulating
CC fragments of the S. pneumoniae genome. Products from the present
CC invention can be used in diagnosis kits and assays, and pharmaceutical
CC compositions and vaccines for S. pneumoniae.
SQ Sequence 28171 BP: 8507 A; 5080 C; 6440 G; 8144 T;

Query Match 17.2%; Score 212.4; DB 1: Length 28171;

Best Local Similarity 53.9%; Pred. No. 5.3e-36;

Matches 533; Conservative 0; Mismatches 436; Indels 19; Gaps 4;

```

QY 61 ATGCGCGTGGGATTCGCGACCGGACCAAAACAAAGATTCGGGTGGCCATACCCCG 120
DB 21711 ATGTAATCGGATTCGCGACCAAAACAAAGATTCGGGTGGCCATACCCCG 21652
QY 121 GCGGCGTGGCGAACTCAACCCGTCGTCATGAGTCTCT--CATCCAGGAGGTGCG 180
DB 21651 GCGGCGTGGCGAACTCAACCCGTCGTCATGAGTCTCT--CATCCAGGAGGTGCG 21592
QY 181 GAGGCGTGGCGAACTCAACCCGTCGTCATGAGTCTCT--CATCCAGGAGGTGCG 240
DB 21541 GAGGCGTGGCGAACTCAACCCGTCGTCATGAGTCTCT--CATCCAGGAGGTGCG 21532
QY 241 GCGGACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB 21531 GCGGACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 21475
QY 301 GAATACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 21474 GAATACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 21415

```

Qy	361	CGTCTGTGCACCCGATGCGTTGTGTGGATTCGCGACACAGCTCAATTGCTTACGAGACCTTC	420
Db	21414	CCAAATATTAGACGATGCTATGTT-----AACGCAAAAAACAACGTAAACCTGTT	21367
Qy	421	CAGACCCCGCCGACCGCGCTACCCCTGCTTGGCCCGCATAGCGAAGTGCCTGCTGACTC	480
Db	21366	CGTACAAATAGAGCAACAACCTACCGCTCTCGTTCCATGATGATGAGTGTGACAGTTGATG	21307
Qy	481	CGCCGCCAGGTTGGCGCTTACCACTGATGCGAACCCAAAGGGGCGCGGTGTCTATG	540
Db	21306	GCTGTTCAAAATCGGAGCTCTCCTTCTTACTTAAGCAAGCTGGGGGCTCGTGTGTCTTACTT	21247
Qy	541	GGCGGGGCGCCGCGCTGCGAACCGCGAGCTGTGTGATGTGGCGCGCGACCGCGCGC	600
Db	21246	GGTGTGTACAGAGTGTTCCTCAAAAGGAAAGTAACTATCATGTGGTGTGTGTCTGCTCGCT	21187
Qy	601	TACAAACGACCCCGCATCTGCCCAAGGCGATGCGCGACCGCTTACGCTGTGACATCAAC	660
Db	21186	ACACATGCTGCCCCCACTGCGCCCTGTGGTGTGGTGTGCTCAAGTACATATTTTAAATTTTGT	21127
Qy	661	ATCGACAATCTTGGCAACTGCAAGCGCGAGTTCTGCGCGCGGATCCACACTGCGTACTCA	720
Db	21126	TCCAAAGGCTCTCTAGTTCTTAAAGAAAGTCTTTGGAAGTCAAAATTCAACTCTTATGTCT	21067
Qy	721	TCGGCTTACGAGCTCGAGGGTGGCCCTCAAACTGTCGCACCTGTGTATTTGGGCGCTCTG	780
Db	21066	AATTCATTCAACATTTAGGAAGAGTGTGAGAGATGCTGAATGTGTATTTGAGACCCATTCTC	21007
Qy	781	GTGCGAGCGCGACAGGACCAACCAATTAGTCTCAATTCCTTGTGCGCATGTGAAACA	840
Db	21006	ATCCCTGTGTGCAAAAGCAGCCGGAATTTGTGTACAGATGATGATGTCAAAACAAATGGCTGCA	20947
Qy	841	GGTGGGTACTGTGTGATATAGCATTCGACACAGGCGCGGTGTTTCA--AGGCTCACGA	897
Db	20946	GGCTCTGTGA--TCGTTTGACCTGTGCTGTGTGACCAAGGTGGCGGTATGAAACAGCTGACCGT	20888
Qy	898	CCGACCACTTACGACGACCCCGAGCTTGGCGGTGCGACGACAGCTGTTTACGTGCGTGGCG	957
Db	20887	GTGACAAACGACGATGAAACCCGCTTATGAAAAACAGGTTGTTCTCACATATGCGGTTCC	20828
Qy	958	AACATGCCCCGCTTGGTGTGCGAAGAGCTGACCTTACGCGCTTACCCACGCGACGATGCCG	1017
Db	20827	AATATCCCTGGTGGGTGTGCTGCGACACTTACACCATGGCCCTTAACCAATGTACTTCTCT	20768
Qy	1018	TATGTGCTCGAGCTTGCACGACATGGCT	1045
Db	20767	TATATCGAAGCTTTGCTGGCAAGGAT	20740
RESULT 7			
V74442/C			
ID	V74442	standard: DNA; 9280 BP.	
AC	V74442:		
DT	16-MAR-1999	(first entry)	
DE	Staphylococcus aureus contig SFQ ID #131.		
FW	Computer readable medium; vaccine; S. aureus infection; immunodetection;		
KW	cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;		
KW	skin infection; surgical wound infection; scalded skin syndrome;		
OS	toxic shock syndrome; ds.		
OS	Staphylococcus aureus.		
FH	Key	Location/Qualifiers	
FT	misc_feature	841..900	
FT		/tag-a	
FT		/note-	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
FT	misc_feature	2641..2700	
FT		/tag-b	
FT		/note-	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
FT			

FT	misc_feature	4441..4500
FT	/tag=	c
FT	/note=	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
FT	/tag=	d
FT	/note=	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
FT	/tag=	e
FT	/note=	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
PN	EP-786519-A2.	
PD	30-JUL-1997.	
PD	07-JAN-1997:	100117.
PR	05-JAN-1996:	US-009861.
PA	(HUMAN) HUMAN GENOME SCI INC.	
P1	Batash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA,	
PI	Rosen CA;	
DR	WPI; 97-374922/35.	
PT	Polynucleotide(s) and proteins derived from Staphylococcus aureus -	
PT	stored on computer readable medium and used in the production of	
PT	anti-S.aureus vaccines	
PS	Claim 1; Page 710-715; 3271bp. English.	
CC	This sequence represents one of 5191 Staphylococcus aureus DNA sequences	
CC	of the invention. The DNA sequences are recorded on a computer readable	
CC	medium, preferably selected from a floppy or hard disk, random access	
CC	memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using	
CC	the S.aureus DNA sequences allows putative functions to be assigned so	
CC	that protein-encoding or regulatory regions of commercial, therapeutic or	
CC	industrial importance can be obtained. Specifically, sequences which are	
CC	likely to encode antigens have been identified and these polypeptides can	
CC	be used in a vaccine composition against S.aureus infection. The	
CC	polypeptides can also be used in a kit for the immunodetection of	
CC	S.aureus in a sample. S.aureus is implicated in numerous human diseases,	
CC	including cellulitis, eyelid infections, food poisoning, osteomyelitis,	
CC	skin and surgical wound infections, scalded skin syndrome, toxic shock	
CC	syndrome, etc. Organisms transformed with the DNA sequences can be used	
CC	for recombinant production of the polypeptides. The new DNA sequences	
CC	(and their fragments) are useful as primers or probes for isolating	
CC	homologues of any of the S.aureus DNA sequences contained on the	
CC	computer readable medium.	
SQ	Sequence 9280 BP; 2958 A; 1549 C; 1309 G; 3060 T;	
<hr/>		
Query Match 14.4%; Score 178.4; DB 1; Length 9280;		
Best Local Similarity 50.6%; Pred.No. 5.8e-29;		
Matches 484; Conservative 0; Mismatches 466; Indels 6; Gaps 2		
DG	151 CATGAGTGTCTATCCAGGCACGTCGCCGAGAGGGCTCGCGATTACGACGGCATTTTC 210	
DB	9280 CATACTGTTTAGTGTGAACAACAATATCGGGTTTGAGATCATTCTTTGAMAATGTAGATTAC 9221	
OY	211 AAGCGGAGAGCGGCCCAACTGCTGCGACACC GCCACAGCTGTGGCCGACCTGTTTA 270	
DB	9220 AAAGAAGCAGGCGTCGAGATTTTCTGTCAACAAGCAAAGTTTGG---CATGTGATATG 9164	
OY	271 TTGCTCAAGGCTAAAGAACCGATAGCGCGGATACGCGCCGCTCGAGACGCGGACATC 330	
DB	9163 GTTATTTAAAGTTAAAGAACCACTTAATCTGAATATTCATATTTTAAAGAAAGGCTTGA 9104	
OY	331 TTGTTACAGCTTCTGCATTGTGGCGCGTCACGTGCTGACCGAGCGGTGTGATTTCC 390	
DB	9103 TTATTACCTTATCTTCATTTTAGCAAAATGACAAAATAATTACACAAAGCTTGTATGATAGA 9044	
OY	391 GGCACACGTCATTGCTACGACAGCGTCCAGACCGCGGACGAGCGGACATACCCCTGCTT 450	

Db 9043 AAGTAATTAGTTCATATGAGACTGTGCACTATACCAAGCAGCTTTACCATTTGTTA 8984
 QY 451 GCCCGATAGAGCAAGTCCGCGGTGCACTGCGCCAGGTTGCGCTTACACATGATG 510
 Db 8983 TCACCAATGAGTAGAGAGAAAGATGTCAAGTTCAGAGAGTTCTTACAA 8924
 QY 511 CGAACCCAGAGGCGCGGTGTGCTGATGCGGGGTGCGCGGTGCAACCGCGGAC 570
 Db 8923 AACTTAATGCTGTATGGAAATCTACTAGTGTGTCTCCAGAGTACCTTAAGGTAAA 8864
 QY 571 GTGCTGATGATGCGCGCGCGGACCGGCTTACAGCAGCGCGCATTCGCAAGCGATG 630
 Db 8863 GTACATATATGCTGTGTGTGTGTCAGAGAGAGCAAAATGCACTTAAATTCAGTAC 8804
 QY 631 GCGCGACCGCTTACGCTTACAGCATCAACATGACAGCAAACTTGGCACTGACGCGGAG 690
 Db 8803 GGTGAGATGTTACGATTTAGATTTAAATCCAAAGCGTTTACACAAATAGATGATTTA 8744
 QY 691 TTGTGGGCGGATTCACACTGCTACTGCTGCTAGAGCTGAGAGGTGCGCTCAA 750
 Db 8743 TTGCTGAGAGGTGATACATCAATATATGTCAAATCCGTTGAATTTGATGTTAAA 8684
 QY 751 CGTGCGACCTGCTGATGATGCGCGCTGCTGCTGCGAGCGCGCAAGCACCATAATGATC 810
 Db 8683 CAATAGTAT.TAGTATAGGTGATGCAATTTAATTCAGG.GCTAAAGCGCAAGACTGTA 8624
 QY 811 TCGAATTCATCTGTCCGCTATGAAACCAAGTGTGCTGCTGCTGATATAGCCATGAC 870
 Db 8623 ACAGAGACATGATTTAAACAAATGAAATGAGTCAATTTATGACATTTGCTATGAT 8564
 QY 871 CAGGCGGCTGTTTGAAGGCTC--ACGACGACACACTAGACAGCAGCGAGCTGCGC 927
 Db 8563 CAGGCGGCTATTTTAAACAACTGATTAATTTAGACACATGATGATCTCATATATAT 8504
 QY 928 GTGACAGACACGCTGTTTACTGCTGCGGCAACATCCGCTGCTGCGGAGAGCTG 987
 Db 8503 AAGCATGATGATGCTTATGATGCAAAATGTCAGAGTGCATACCGGCTACTTGG 8444
 QY 988 ACTTACGCTGACCAACCGAGATGCGCTATGCTGATGCTGCTGCTGCTGCTGCTG 1047
 Db 8443 AGCTTAGCTTTAAATATCTACGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 8384
 QY 1048 CCGGCGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1103
 Db 8383 AGAGAGCATTTAAATCAATCAACATATATATATAGTTTAAATCTTACAAAG 8328
 RESULT 8
 T67971
 ID T67971 standard; DNA; 1074 BP.
 AC T67971;
 DT 15-JUL-1997 (first entry)
 DE H. pylori membrane protein ORF 05cp20518orf1.
 KW Vaccine; prevention; treatment; infection; identification;
 KW binding compound; bacterium; life cycle; activator; bacteria;
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis;
 KW membrane; amino acid; metabolism; ds.
 OS Helicobacter pylori.
 FH key Location/Qualifiers
 FT 1..1074
 FT /tag="a
 FT /note="no stop codon given"
 PN W09640893-AL.
 PD 19-DEC-1996.
 PF 06-JUN-1996; U091122.
 PR 07-JUN-1995; US-487032.
 PR 01-APR-1996; US-630405.
 PA (ASTR) ASTRA AB.
 PI Berghindh OT, Smith D, Mellgaard BL.
 DR WPI: 97-052306/05.
 PT Helicobacter pylori nucleic acid sequences and related
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori

PT Infection, and to detect Helicobacter
 PS Claim 27; Page 819; 148bp; English.
 CC The present sequence encodes a Helicobacter pylori membrane
 CC protein likely to contain four membrane spanning regions.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds,
 CC useful as potential H. pylori life cycle activators or inhibitors.
 CC The genomic sequence of H. pylori (ATCC 55679) was determined from
 CC overlapping contigs generated by mechanically shearing the
 CC bacterial DNA. The sequences were analysed for ORF of at least 180
 CC nucleotides, and the predicted coding regions defined by computer
 CC evaluation. To identify likely H. pylori antigens for vaccine
 CC development, the amino acid sequences predicted from various ORF
 CC were analysed for significant homology to other known or exported
 CC membrane proteins. Having identified and determined the sequences
 CC of interest, particular regions can be isolated from H. pylori by
 CC PCR amplification for recombinant polypeptide production, e.g. in
 CC E. coli hosts.
 SQ Sequence 1074 BP; 302 A; 189 C; 297 G; 286 T;

Query Match 7/7% Score 94.8; DB 1; Length 1074;

Best Local Similarity 46.1%; Pred. No. 1.2e-11; Matches 370; Conservative 0; Mismatches 417; Indels 15; Gaps 1;

QY 265 GATTATGCTCAAGTCAAGAACGATAGCGGGAATAGCGCGCTGCGACAGCGG 324
 Db 133 GATTGTGCTCAAAATGCAAGAGCCTTAGACATGATATCCCTTCTCAAAAGAAA 192
 QY 325 CAGATCTGTACGCTTCTTGCAATTTGCGCGCTGACGCTGCTGACCGATGCTGTTG 384
 Db 193 GCGACTGCTGTTAGTATTTGATTTAGCTATCAAAAAAGCTGTGGAATGTTAT 252
 QY 385 GATTCGCGACACGCTCAATGCTTACGAGACCGGCGAGCGCGGCGGCGACTACCC 444
 Db 253 AATTAATAAATATCTTCTTATTTGCACTAATCAATGCGCGGCTTAAAGAACTACCT 312
 QY 445 CTGCTGCGCGGATGAGGAAAGTCCGCTGCTGCTGCGCGCGGCTGCTGCTGCTGCTG 496
 Db 313 ATTTAGCGGCTTATGAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 372
 QY 496 -----GCTTACCACTGATGCAACCAAGGCGCGCGCTGCTGCTGCTGCTGCTGCT 549
 Db 373 TTACTGCTTATGACATGTTAAAGGCTTTATGGTTAAAGGCGCTGCTGCTGCTGCTG 432
 QY 550 CCGGCGCTGCAACCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 609
 Db 433 TCGGCTGCGCAAGAGCGCTAAATATCTCTTATTTGAGCGCGCTGCTGCTGCTGCTG 492
 QY 610 GCGCGCATGCGCAACGCGATGCGCGCGCTTACGCTTACGATCAACATGACAGAAA 669
 Db 493 GCGAAAGCTTAAAGCAATAGGCGCGCTTAAAGTAAAGTAAAGTAAAGTAAAGTAA 552
 QY 670 CTTCGCACTGAGCGGAGTGTGCGCGCGGATCCACATGCTACTATCGGCTAC 729
 Db 553 TTACAAACACCTTATATATATTTATATATTTAGAGCTTAAAGCTTAAAGTAAAG 612
 QY 730 GAGCTGAGAGGTCGCTCAAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 789
 Db 613 AATATCATTCAGGCTTAAAGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 672
 QY 790 GCGAAGGACCCAAATTTGCTGCAATTCCTTGTGCGGCTTAAAGAAACGAGTGGGTA 849
 Db 673 AGCCAAACCCCTTAAATGATATTTAAAGAGCATTTAAATATCATGACAGCAAGGGTA 732
 QY 850 CTGCTGATATAGCCATGACAGCGCGCTGCTTGAAGGCTGACAGCAGCAGCAGCTAC 909
 Db 733 GTATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 792
 QY 910 GACCAACCGAGTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 969
 Db 793 TCTAACCGGCTGATGAGAAAGTTGTTGATTTATGCTGCTGCTGCTGCTGCTGCTGCT 852

FT	misc_feature	19201..19260	
FT	/tag=	k	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
FT	/note=		
FT	misc_feature	21001..21060	
FT	/tag=	l	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
FT	/note=		
FT	misc_feature	22801..22860	
FT	/tag=	m	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
FT	/note=		
FT	misc_feature	24601..24660	
FT	/tag=	n	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
FT	/note=		
FT	misc_feature	26401..26460	
FT	/tag=	o	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
FT	/note=		
FT	misc_feature	28201..28260	
FT	/tag=	p	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
FT	/note=		
FT	misc_feature	30001..30060	
FT	/tag=	q	"these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"
FT	/note=		
FT	misc_feature		
PN	Ep-786519-A2.		
PD	30-JUL-1997.		
PF	07-JAN-1997.	100117.	
PR	05-JAN-1996:	US-009861.	
PA	(HUMA-) HUMAN GENOME SCI INC.		
PI	Batash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA,		
PI	Rosen CA;		
DR	WPI: 97-374922/35.		
PJ	Polynucleotide(s) and proteins derived from Staphylococcus aureus -		
PT	stored on computer readable medium and used in the production of		
PS	anti-S.aureus vaccines		
CS	Claim 1: Page 452-469: 327pp; English.		
CC	This sequence represents one of 5101 Staphylococcus aureus DNA sequences of the invention. The DNA sequences are recorded on a computer readable medium, preferably selected from a floppy or hard disk, random access memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using the S.aureus DNA sequences allows putative functions to be assigned so that protein-encoding or regulatory regions of commercial, therapeutic or industrial importance can be obtained. Specifically, sequences which are likely to encode antigens have been identified and these polypeptides can be used in a vaccine composition against S.aureus infection. The polypeptides can also be used in a kit for the immunodetection of S.aureus in a sample. S.aureus is implicated in numerous human diseases, including cellulitis, eyelid infections, food poisoning, osteomyelitis, skin and surgical wound infections, scaled skin syndrome, toxic shock syndrome, etc. Organisms transformed with the DNA sequences can be used for recombinant production of the polypeptides. The new DNA sequences (and their fragments) are useful as primers or probes for isolating homologues of any of the S.aureus DNA sequences contained on the computer readable medium.		
SQ	Sequence 31096 BP;	11857 A;	5243 C; 5477 G; 7488 T;

[illegible]

RESULT 11
Q21833/c
ID Q21833 standard; DNA: 390 BP.
AC Q21833;
DE 08-JUN-1992 (first entry)
KM Randomising oligonucleotide used in SPERT mRNA prep.
DE Systematic polypeptide evolution by reverse translation: SPERT;
KM ligand binding: ss.
OS Synthetic.
PN MO9202535-A.
PD 20-FEB-1992.
PF 01-AUG-1991: US-561968.
PR 02-AUG-1990: US-561968.
PA (COLS) UNITV OF COLORADO.
PI Gold L, Tuerk C.
DR WPI: 92-080018/10.
PT New method of systematic polypeptide evolution by reverse translation - by linking each polypeptide in sample mixt. to individualised mRNA allowing further synthesis of selected polypeptide(s)
PS Example: Page 55; 102pp; English.
CC The sequence is that of an example randomising oligonucleotide which is used in the prep. of mRNA encoding candidate polypeptides for the method of systematic polypeptide evolution by reverse translation (SPERT). The method provides a rapid way of isolating and identifying polypeptide ligands which bind to target mols. The polypeptide ligands can be used in e.g. assay methods, diagnostic procedures, cell sorting, as activators or inhibitors of target mol. function, as probes, as sequestering agents, drug delivery vehicles, modifiers of hormone action and as catalysts. See also Q21830-Q21832.
CC Sequence 390 BP; 125 A; 126 C; 133 G; 6 T;
SQ

Query Match 4.1%; Score 50.8; DB 1; Length 390;
Best Local Similarity 46.2%; Pred. No. 0.016;
Matches 169; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 426 CGCGAGCGGCGCACTACCCCTGCTTGGCCCGATGAGGAGTGGCGCGTGCAGTCCGC 485
DB 366 CGTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 307
QY 486 CCAGGTTGGCGCTTACCACTGATGCAACCCAGGGGCGCGGTGCTGATGGCGG 545
DB 306 CGTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 247
QY 546 GGTGCGCGGCGTGAACCGCGCGAGTGTGATGGCGCGCGCGCGCGCTACAA 605
DB 246 CGTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 187
QY 606 CGCAGCGCGCATGCGCAAGCGATGAGCGGCGGCGGATGCTTACATCAACATGCA 665
DB 186 CGTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 127
QY 666 CAAACTTGGCACTGAGCGCGGAGTGTGCGCGCGGATCCACACTGCTACTCATCGC 725
DB 126 CGTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 67
QY 726 CTACGAGCTGAGAGGTCCCTCAAACTGCGGAGCTGTGTTGGGGCGCTCTGTGCGC 785
DB 66 CGTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7
QY 786 AGGCGC 791
DB 6 TGGCCC 1

RESULT 12
Q36859/c
ID Q36859 standard; DNA: 390 BP.
AC Q36859;
DE 22-JUN-1993 (first entry)

DE PCR primer for 5' fixed sequence contg. T7 promoter and RBS.
KM Systematic peptide evolution by reverse translation: SPERT; ligand;
KM specific: inhibitors; probes; assay; cell sorting; ss.
OS Synthetic.
PN MO9303172-A.
PD 18-FEB-1993.
PF 31-JAN-1992: U00801.
PR 01-AUG-1991: US-739055.
PA (UTRE-) UNITV RES CORP.
PI Gold L, Prihnow D, Smith JD, Tuerk C;
DR WPI: 93-076529/09.
PT Systematic polypeptide evolution by reverse translation - used for prodn. of polypeptide ligand specific for desired target molecule
PS Example 1; Page 84; 98pp; English.
CC SPERT is used to select novel polypeptides that bind the antibody of the epitope commonly recognised by the antisera from autoimmune mice which are the f1 progeny of a cross of NZB and NZW parents (Portanova et al., J. Immunol. 144, 4633, 1990). The known epitope consists of ca. 10 amino acids at the N-terminus of the histone H2B protein. To make mRNA encoding candidate polypeptides a 5' fixed sequence composed of a T7 promoter sequence and a ribosome binding site which is recognised by both prokaryotic and eukaryotic ribosomes, terminating in a restriction endonuclease site is synthesised and cloned using a number of oligonucleotides (example shown). A 3' fixed sequence is placed into a restriction site to provide an mRNA encoding the C-terminal trailer sequence of ca. 100 nucleotides lacking stop codons. In addition, a 3' primer annealing site is provided so that cDNA syntheses can be accomplished on the mRNA recovered from partitioned CC ribosome complexes. See also Q36845-63.
CC Sequence 390 BP; 125 A; 126 C; 133 G; 6 T;
SQ

Query Match 4.1%; Score 50.8; DB 1; Length 390;
Best Local Similarity 46.2%; Pred. No. 0.016;
Matches 169; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 426 CGCGAGCGGCGCACTACCCCTGCTTGGCCCGATGAGGAGTGGCGCGTGCAGTCCGC 485
DB 366 CGTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 307
QY 486 CCAGGTTGGCGCTTACCACTGATGCAACCCAGGGGCGCGGTGCTGATGGCGG 545
DB 306 CGTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 247
QY 546 GGTGCGCGGCGTGAACCGCGCGAGTGTGATGGCGCGCGCGCGCTACAA 605
DB 246 CGTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 187
QY 606 CGCAGCGCGCATGCGCAAGCGATGAGCGGCGGCGGATGCTTACATCAACATGCA 665
DB 186 CGTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 127
QY 666 CAAACTTGGCACTGAGCGCGGAGTGTGCGCGCGGATCCACACTGCTACTCATCGC 725
DB 126 CGTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 67
QY 726 CTACGAGCTGAGAGGTCCCTCAAACTGCGGAGCTGTGTTGGGGCGCTCTGTGCGC 785
DB 66 CGTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7
QY 786 AGGCGC 791
DB 6 TGGCCC 1

RESULT 13
X34205
ID X34205 standard; DNA: 1413 BP.
AC X34205;
DE 06-JUL-1999 (first entry)
DE Mycobacterium species nucleic acid sequence 415.
KM Secreted protein; Mycobacterium; primer; PCR; amplification; probe;

KW hybridisation; detection; vaccine; immunisation; infection; ss.
 OS Mycobacterium sp.
 PN WO9909186-A2.
 PD 25-FEB-1999.
 PF 14-AUG-1998; F01813.
 PR 11-SEP-1997; FR-011325.
 PR 14-AUG-1997; FR-010404.
 PA (INSP) INST PASTEUR.
 PI Guigueno A;
 DR WPI: 99-181045/15.
 DR P-PSDB: Y04952.
 PT Mycobacterial DNA vectors containing reporter constructs - for
 PT identifying coding or promoter sequences involved in
 PT infection-associated protein expression
 PS Claim 22; Fig 41S: 309pp; French.
 CC Sequences X34001-X34252 represent nucleic acids encoding secreted
 CC proteins from various Mycobacterium species microorganisms. The
 CC nucleotide sequences can be used as primers and probes for methods
 CC for detecting and identifying mycobacteria, especially belonging to
 CC the M. tuberculosis complex. The encoded proteins can be used in
 CC vaccines for immunisation against a bacterial or viral infection.
 SO Sequence 1413 BP: 187 A: 459 C: 540 G: 227 T:

Query Match 4.1%; Score 50.4; DB 1; Length 1413;
 Best Local Similarity 43.0%; Pred. No. 0.021;
 Matches 246; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

QY 69 CGGATTCGAGACGAGAACCAACAAACGATTCGGGTGGCCATCACCCCGCGCGGT 128
 Db CGGATTCGAGACGAGAACCAACAAACGATTCGGGTGGCCATCACCCCGCGCGGT 128
 QY 129 CGCGAATCAACCCGCTGTCGATGAGTGTCTCATCCAGGAGTGGCGGAGAGGCTC 188
 Db GCGCGGCTGTGGCGCGCGCTGTCGATGAGTGTCTCATCCAGGAGTGGCGGAGAGGCTC 188
 QY 189 GGTATCAACGAGCGGATTTCAAGGCGGAGCGCGCACTGTGCGACCGCGACCA 248
 Db GGAACATATCTCTGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGA 248
 QY 249 GGTGTGGCGCGCGCTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 308
 Db GGGGTGTGGCGCGCGCTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 308
 QY 249 GGGGTGTGGCGCGCGCTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 308
 Db GGGGTGTGGCGCGCGCTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 308
 QY 309 CCGGCTGCGACGAGCGGAGATCTTGTTCACGTTCTTGTTCATTTGGCGCGGTACGTCGCTG 368
 Db CCGGCTGCGACGAGCGGAGATCTTGTTCACGTTCTTGTTCATTTGGCGCGGTACGTCGCTG 368
 QY 309 CCGGCTGCGACGAGCGGAGATCTTGTTCACGTTCTTGTTCATTTGGCGCGGTACGTCGCTG 368
 Db CCGGCTGCGACGAGCGGAGATCTTGTTCACGTTCTTGTTCATTTGGCGCGGTACGTCGCTG 368
 QY 369 CACCGATGCGTTGTGGATTCGCGACGACGTCATTTGCTACGAGACCGTCCAGACCGC 428
 Db GCTGTGTTCTGTGGAGCGGTGCGCTCCACCATCTGCGGGGACGACCTGCGCGGGGGGT 428
 QY 429 CGAGGCGGCACTACCCCTGCTGTGCGCGGATGAGGAGTGGCGGTGACTGCGCGCCCA 488
 Db CTACGAACCTTCGCGCGCGGTGCGCGGACGACGAGGCTACCAAGCGCGCGCGCGCA 488
 QY 489 GGTGGCGCTTACCACTGATGCGAACCAAGGCGCGCGGTGCTGATGAGGCGGGGT 548
 Db ATTGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 548
 QY 549 GCGCGGCTGCAACCGGCGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 608
 Db TCGACTCAAGGAATTTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 608
 QY 609 AGCCGCAATCGCAACGCGCATGGCGGCGACCG 640
 Db CGTACAGCTCCACACCGACGACGCGCGGTGCG 640

RESULT 14
 X34206
 ID X34206 standard; DNA; 1722 BP.

AC X34206;
 DE 06-JUL-1999 (first entry)
 DT Mycobacterium species nucleic acid sequence 41T.
 KW Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
 KW hybridisation; detection; vaccine; immunisation; infection; ss.
 OS Mycobacterium sp.
 PN WO9909186-A2.
 PD 25-FEB-1999.
 PF 14-AUG-1998; F01813.
 PR 11-SEP-1997; FR-011325.
 PR 14-AUG-1997; FR-010404.
 PA (INSP) INST PASTEUR.
 PI Guigueno A;
 DR WPI: 99-181045/15.
 DR P-PSDB: Y04953.
 PT Mycobacterial DNA vectors containing reporter constructs - for
 PT identifying coding or promoter sequences involved in
 PT infection-associated protein expression
 PS Claim 22; Fig 41T: 309pp; French.
 CC Sequences X34001-X34252 represent nucleic acids encoding secreted
 CC proteins from various Mycobacterium species microorganisms. The
 CC nucleotide sequences can be used as primers and probes for methods
 CC for detecting and identifying mycobacteria, especially belonging to
 CC the M. tuberculosis complex. The encoded proteins can be used in
 CC vaccines for immunisation against a bacterial or viral infection.
 SO Sequence 1722 BP: 231 A: 548 C: 657 G: 286 T:

Query Match 4.1%; Score 50.4; DB 1; Length 1722;
 Best Local Similarity 43.0%; Pred. No. 0.021;
 Matches 246; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

QY 69 CGGATTCGAGACGAGAACCAACAAACGATTCGGGTGGCCATCACCCCGCGCGGT 128
 Db CGGATTCGAGACGAGAACCAACAAACGATTCGGGTGGCCATCACCCCGCGCGGT 128
 QY 378 GCGGACTGCGCGCGCGCGCTGTCGCGGCTATTCGCGCGCTGAGCGCAACCCCTCGG 437
 Db GCGGACTGCGCGCGCGCGCTGTCGCGGCTATTCGCGCGCTGAGCGCAACCCCTCGG 437
 QY 129 CGCGAATCAACCCGCTGTCGATGAGTGTCTCATCCAGGAGTGGCGGAGAGGCTC 188
 Db GCGCGGCTGTGGCGCGCGCTGTCGATGAGTGTCTCATCCAGGAGTGGCGGAGAGGCTC 188
 QY 438 GCGCGGCTGTGGCGCGCGCTGTCGATGAGTGTCTCATCCAGGAGTGGCGGAGAGGCTC 497
 Db GCGCGGCTGTGGCGCGCGCTGTCGATGAGTGTCTCATCCAGGAGTGGCGGAGAGGCTC 497
 QY 189 GGTATCAACGAGCGGATTTCAAGGCGGAGCGCGCACTGTGCGACCGCGACCA 248
 Db GGAACATATCTCTGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGA 248
 QY 498 GGAACATATCTCTGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGA 557
 Db GGAACATATCTCTGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGA 557
 QY 249 GGTGTGGCGCGCGCTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 308
 Db GGGGTGTGGCGCGCGCTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 308
 QY 558 GGGGTGTGGCGCGCGCTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 308
 Db GGGGTGTGGCGCGCGCTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 308
 QY 309 CCGGCTGCGACGAGCGGAGATCTTGTTCACGTTCTTGTTCATTTGGCGCGGTACGTCGCTG 368
 Db CCGGCTGCGACGAGCGGAGATCTTGTTCACGTTCTTGTTCATTTGGCGCGGTACGTCGCTG 368
 QY 618 CCGGCTGCGACGAGCGGAGATCTTGTTCACGTTCTTGTTCATTTGGCGCGGTACGTCGCTG 677
 Db CCGGCTGCGACGAGCGGAGATCTTGTTCACGTTCTTGTTCATTTGGCGCGGTACGTCGCTG 677
 QY 369 CACCGATGCGTTGTGGATTCGCGACGACGTCATTTGCTACGAGACCGTCCAGACCGC 428
 Db GCTGTGTTCTGTGGAGCGGTGCGCTCCACCATCTGCGGGGACGACCTGCGCGGGGGGT 428
 QY 678 GCTGTGTTCTGTGGAGCGGTGCGCTCCACCATCTGCGGGGACGACCTGCGCGGGGGGT 737
 Db GCTGTGTTCTGTGGAGCGGTGCGCTCCACCATCTGCGGGGACGACCTGCGCGGGGGGT 737
 QY 429 CGAGGCGGCACTACCCCTGCTGTGCGCGGATGAGGAGTGGCGGTGACTGCGCGCCCA 488
 Db CTACGAACCTTCGCGCGCGGTGCGCGGACGACGAGGCTACCAAGCGCGCGCGCGCA 488
 QY 738 CTACGAACCTTCGCGCGCGGTGCGCGGACGACGAGGCTACCAAGCGCGCGCGCGCA 797
 Db CTACGAACCTTCGCGCGCGGTGCGCGGACGACGAGGCTACCAAGCGCGCGCGCGCA 797
 QY 489 GGTGGCGCTTACCACTGATGCGAACCAAGGCGCGCGGTGCTGATGAGGCGGGGT 548
 Db ATTGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 548
 QY 798 ATTGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 857
 Db ATTGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 857
 QY 549 GCGCGGCTGCAACCGGCGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 608
 Db TCGACTCAAGGAATTTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 608
 QY 858 TCGACTCAAGGAATTTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 917
 Db TCGACTCAAGGAATTTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 917
 QY 609 AGCCGCAATCGCAACGCGCATGGCGGCGACCG 640
 Db CGTACAGCTCCACACCGACGACGCGCGGTGCG 640
 QY 918 CGTACAGCTCCACACCGACGACGCGCGGTGCG 949

RESULT 15

T68715
T68715 standard; DNA: 15872 BP.

DE Streptomyces venezuelae polyketide synthase vep ORF1.
KM Polyketide synthase; polyhydroxyalkanoate monomer synthase;
OS Streptomyces venezuelae.

Metabolic engineering: ss.

Location/Qualifiers

cds 20..13912

cds 14056..14136

cds 14148..15827

cds W09722711-A1.

26-JUN-1997.

18-DEC-1996; U20119.

19-DEC-1995; U6-008847.

(MINU) UNIV MINNESOTA.

Sherman DR, Williams MD, Xue Y;

WPI: 97-341701/31.

P-PSDB: W19629-30 AND W00918.

Expression cassettes for production of polyhydroxyalkanoate(s) -

provide wide range of biodegradable polymers for medical or

industrial use

Claim 54; Fig 23; 91pp; English.

Streptomyces venezuelae vep ORF1 (T68715) comprises the polyketide

synthase (PKS) gene cluster encoding a polyene of 12 carbons (see

also W19629-30 and W00918). It contains 5 PKS modules, with a 5'

loading module and a 3' end domain. Each of the sequenced modules

includes a keto-ACP, an acyltransferase, a dehydratase, a keto-

reductase and an acyl carrier protein domain. The gene cluster was

cloned using a heterologous hybridisation strategy from a genomic

DNA library. A novel expression cassette encoding the first module

from the vep gene cluster and module 7 from the Streptomyces tyli

gene cluster has polyhydroxyalkanoate (PHA) monomer synthase

activity and can be used for PHA prodn. in host (esp. insect) cells

for use as a biodegradable polymer.

Sequence 15872 BP; 2093 A; 6313 C; 5503 G; 1963 T;

Query Match 4.18; Score 50.4; DB 1; Length 15872;

Best Local Similarity 47.78; Pred. No. 0.023;

Matches 187; Conservative 0; Mismatches 196; Indels 9; Gaps 1;

QY 417 CGTCAGACCCGCGACGACCTGCTGCCCCGATGAGCGAGTCGCCGTGC 476

Db 1540 CCTCCTGACGGGCTGCGGCGCCCTGCGCGGAGCGCCCGCGCGTGTACCCG 1599

QY 477 ACTGCGCGCCGAGTTGGCGCTTACACATGATGCAACCCAGAGGGCGCGGTGCT 536

Db 1600 CACCCCG 1659

QY 537 GATGGCGGGGTGCG 596

Db 1660 CATGGGCTGAGTTGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1719

QY 597 CGGCTACACGACCG 649

Db 1720 CGCGGACGTGACCG 1779

QY 649 -CTAGACATCAACATCGCAACTTCGCAACTGAGCGCGCGCGCGCGCGCGCGCG 707

Db 1780 CTTGACCG 1839

QY 708 CACTCGCTACTCATCGGCTGAGAGTGTGCGTCAACGTCGCGCGCGCGCGCGCGAT 767

Db 1840 CCTGCTGAGTCTCTGGGGCGTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGAT 1899

QY 768 TGGGCGCGCTCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 799

Db 1900 CAGCG 1931

Search completed: June 22, 2000, 15:07:26
Job time: 17772 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 15:03:42 ; Search time 341.15 Seconds
(without alignments)
471.322 Million cell updates/sec

Title: US-09-362-485-4
1237

Perfect score: 1 ATCTGCGATGATGCGAC.....GGAAGCGATGATGCGCCG 1237

Sequence: IDENTITY_NOC

Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

Issued Patents_NA:*
1: /cgn2.6/ptodata/2/ina/5A.CONB.seq:*
2: /cgn2.6/ptodata/2/ina/5B.CONB.seq:*
3: /cgn2.6/ptodata/2/ina/5C.CONB.seq:*
4: /cgn2.6/ptodata/2/ina/5D.CONB.seq:*
5: /cgn2.6/ptodata/2/ina/6.CONB.seq:*
6: /cgn2.6/ptodata/2/ina/PTCUS.CONB.seq:*
7: /cgn2.6/ptodata/2/ina/Backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47.5	3.8	1620	3	US-08-461-775-10
2	47.6	3.8	28958	1	US-08-258-261B-6
3	47.6	3.8	28958	1	US-08-456-837-6
4	47.6	3.8	28958	1	US-08-457-342-6
5	47.6	3.8	28958	1	US-08-457-646A-6
6	47.6	3.8	28958	2	US-08-458-076A-6
7	47.6	3.8	28958	2	US-08-764-233A-4
8	47.6	3.8	28958	2	US-08-457-335A-6
9	47.6	3.8	28958	2	US-08-729-214-6
10	47.6	3.8	28958	2	US-08-764-233A-1
11	47.6	3.7	2582	4	US-08-816-105A-2
12	45.4	3.7	833	3	US-08-403-852D-3
13	45.4	3.7	5392	3	US-08-403-852D-6
14	45.4	3.4	1833	3	US-08-403-852D-6
15	45.4	3.4	2185	1	US-08-177-508-3
16	45.4	3.4	2185	1	US-08-263-310-3
17	45.4	3.3	2414	7	5248599-1
18	45.4	3.3	2668	3	US-08-461-775-11
19	41	3.3	20235	2	US-07-642-734C-3
20	40.8	3.3	44377	4	US-08-804-227C-7
21	40.8	3.3	44377	4	US-08-804-198-1
22	40.6	3.3	4257	3	US-08-690-473-1
23	40.6	3.3	12001	2	US-08-458-568A-11
24	40.2	3.2	2064	1	US-08-343-428-1
25	40.2	3.2	11219	2	US-07-642-734C-1
26	39.6	3.2	474	3	US-08-403-852D-14
27	39.6	3.2	1524	7	5512669-1

28	39.6	3.2	1525	7	5229279-1	Patent No. 5229279
29	39.4	3.2	459	4	US-08-387-942C-35	Sequence 35, Appl
30	39.4	3.2	1998	4	5212296-8	Patent No. 5212296
31	39.4	3.2	12588	4	US-08-387-942C-1	Sequence 1, Appl
32	39.2	3.2	420	1	US-08-470-179-148	Sequence 18, App
33	39.2	3.2	2353	6	PCT-US92-06840-1	Sequence 1, Appl
34	39	3.2	43280	3	US-08-804-227C-1	Sequence 1, Appl
35	38.8	3.1	8051	4	US-08-576-626A-2	Sequence 2, Appl
36	38.8	3.1	11219	2	US-07-642-734C-1	Sequence 1, Appl
37	38.6	3.1	1187	2	US-08-440-856A-2	Sequence 2, Appl
38	38.6	3.1	3231	2	US-08-074-121-4	Sequence 4, Appl
39	38.6	3.1	3231	6	PCT-US94-08447-4	Sequence 4, Appl
40	38.4	3.1	1215	4	US-08-947-726A-1	Sequence 1, Appl
41	38.4	3.1	2048	5	US-08-776-251-1	Sequence 1, Appl
42	38.2	3.1	2109	4	US-08-555-568B-20	Sequence 20, Appl
43	38	3.1	30001	1	US-08-125-458-1	Sequence 1, Appl
44	38	3.1	30001	3	US-08-474-933-1	Sequence 1, Appl
45	37.8	3.1	2588	3	US-08-796-414B-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-461-775-10
Sequence 10, Application US/08461775
Patent No. 5858773
GENERAL INFORMATION:
APPLICANT: MAZODIER, Philippe
APPLICANT: GUGLIEM, Gerard
TITLE OF INVENTION: REGULATORY NUCLEOTIDE SEQUENCE OF THE
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESS: Burns, Doane, Swecker & Mathis
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22315-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461.775
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/050.313
FILING DATE: 10-MAY-1993
APPLICATION NUMBER: FR 9011186
FILING DATE: 10-SEP-1990
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feury, Sharon E
REGISTRATION NUMBER: 36,113
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1620 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1620
US-08-461-775-10

Query Match 3.8%; Score 47.6; DB 3; Length 1620;
 Best Local Similarity 45.7%; Pred. No. 0.003;
 Matches 321; Conservative 0; Mismatches 369; Indels 12; Gaps 4;

```

QY 114 CACCCGCGCGCGCTGCGGAGTAAACCGTGTGGCCATGAGTCTCATCCAGGAGG 173
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 879 CACCTTCACCGCTGACACCGTATCCGAGAGAGTGGCTCAAGCTCGACAGGCGG 938
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 174 TGCCGAGAGAGGCTCGGTATCACCAGCGGATTTCAAGGCGGAGCGGCGCAACTGCT 233
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 939 TCTGAGCTGTGGGACACCGCGCGCGGTACCGTCAACAGGACGACAGCAGCTGCT 998
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 234 CCGGACCGCGCGCGAGGTGTGGCGGAGCTGATTTATTTGCTCAAGCTCAAGAGAGCAGT 293
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 999 GGAGGGGGGGGCGCAAGCGCGAGAGCTCCAGAGGGCGGCTGCCCGCAAGTCAAGAGGCGAGAT 1098
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 294 AGCGGCGGATATACGGCGCGCTGCGACAGCGGAGATCTTTGTCACGTTCTTGCAATTGGC 353
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1059 CGAGTCCAGCCAGCTCGAGCTGGAGCCGAGAGAGCTCCAGAGGCGCTCGCAAGCTGCGC 1118
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 354 C-GGCTCAGTGTGCTGACCGATGGGTGTGGATTCCGGCAGCCAGCTCAATTGCTACG 412
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1119 CGCGGCGCTGTGCTGTGATCCCGTGGCGCGCGCCAGAGTCCAGTCAAGAGAGCGCA 1178
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 413 AGACCGTCCAGACCGCGCGAGCGGAGTACCCCTGCTTGCCTGATGAGCGAA--GTGCG 470
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1179 GCACCGTCTGGAGAGAGCGCATCTCCGAGACCGCGCGCGCTGACAGAGGAGGCGATGCTC 1238
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 471 CGGTGACCTCCCGCGCGAGTGGCGCTTACCA---CCTGATGGGAAGCCAAAGGCGCG 527
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1239 CGGTGTGTGCTCCCGCTGCTGACCGGTCAAGTCTTGAGCAGCAACCTTGCGCGAC 1298
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 528 CGGTGTGTGATGGGCGGTGTGCGCGTGAACCGCGCAGCTGTGTGATGCG---585
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1299 CGCGGAGAGCGCGACCGGTGTGCGGTGTGCTCCGCGCGCGCTCGAGCGCGCTGCGTG 1358
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 585 ---CGCGGCGACCGCGGTCAACAGCAGCGCGCATGCGCAAGCGGATGGCGCGACCGT 641
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1359 GATGCGCGAGAGCGCGCGCTGAGGAGCTACGTCACACCAAGGCGCGGAGCTGGA 1418
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 642 TACGTTCTACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 701
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1419 CAAGGCGCGAGGCTTCAACCGCGCGACCGCGAGTACGCGCACTGTGCAAGCGCGCGT 1478
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 702 GATCCACACTGCTACTATCATGCGCTTACAGCTGAGGAGGTGCGCTCAACAGTGGCGACT 761
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1479 CATGAGCGCGGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1538
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 762 GGTGATTGGGCGCTCTGTGTGCCAGGCGCGCAAGCGCACCA 803
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1539 GCTCTGACGACGAGACCTGTGTGTGAGAGAGCGCGCGCA 1580
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 2
 US-08-258-261B-6
 ; Sequence 6, Application US/08258261B
 ; Patent No. 5639949
 ; GENERAL INFORMATION:
 ; APPLICANT: Schupp, Thomas
 ; APPLICANT: Ligon, James M.
 ; APPLICANT: Beck, James Joseph
 ; APPLICANT: Hill, Dwight Steven
 ; APPLICANT: Ryals, John Andrew
 ; APPLICANT: Gaffney, Thomas Deane
 ; APPLICANT: Lam, Stephen Ting
 ; APPLICANT: Hammer, Phillip E.
 ; APPLICANT: Uknes, Scott Joseph
 ; TITLE OF INVENTION: Genes for the syntheses of
 ; TITLE OF INVENTION: antipathogenic substances
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Ciba-Geigy Corporation

```

STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/258,261B
FILING DATE: 08-JUN-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHEtical: NO
ANTI-SENSE: NO
US-08-258-261B-6

```

Query Match 8%; Score 47.6; DB 1; Length 28958;
 Best Local Similarity 44.1%; Pred. No. 0.0074;
 Matches 286; Conservative 0; Mismatches 359; Indels 3; Gaps 2;

```

QY 148 GGCATAGAGTGTCTATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 207
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25306 GCCCAGCAGCACTCACCTCCGCGAGTGTGCTTATTCGCTGCGCACACCGCGCGCAC 25365
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 208 TTCAGGCGGCGAGCGCGGCGTGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGT 267
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25366 TTCAGGCGGCGCGCGCGTGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 25425
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 268 TTATGCTCAAGTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 327
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25426 TCGCTCCCGCGAGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 25485
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 328 ATCTGTTCACGTTCTTGTGATTTGGCGCGCTCACGTCCTTGACACCGATGCTGTGTGAT 387
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25486 AAGCTGTCTTGTCTTCTTGTGCGCAAGGCTTCAGTGGAGAGGATGCGCTCTGTGCTG 25545
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 388 TCCGCGACCACTCAATTTGCTTACGAGACCGTCCAGACCGCGCGCGCGCGCGCGCGCTG 447
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25546 CTGACCTCTCGCGCGCTTCTCGCGCTCACCTGGAAGATGCGAGCGCGCGCTGCTCT 25605
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 448 CTGCGCGCGATGAGCGAGTCCGCGGTGACCTGCGCGCGCGCGCGCGCGCGCGCGCTTAC 505
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25606 CAGCTCAGTGTGAGCGCTGCGCGGTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 25665
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 506 TGATGCGCAACCGCAAGGCGCGCGCGGTGTGATGGCGGGGTGCCCGGCTGGAACCG 565
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25666 CGGCTCAGCTGTACAGCGCGCGCTTTTCCGTATGTCTTCTGTGGGCGCTCTGTG 25725
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 566 CCGAGC-TGCTGTGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 624
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25726 CGCTCGCTCGCGGTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 25785
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 625 GGCATGGGCGCGCGCTTACGTTTACATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGAT 684
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 25786 GCCTTCGTCAGAGCGCTCTCTCCCTCAGAGAGCGCGCCGCAATCGCCGCTCCGAGC 25845
QY 685 GCGGAGTTCGCGCGCGGATCCACATCGCTACTCATCGGCTTACAGAGCTCGAGAGTGGC 744
Db 25846 AAGCGCTACACACCGTGCCTGCGGCAAGGAGGCGCATGGCGCGCTGCGAGCTCGGCGCTCC 25905
QY 745 GTCAACGTCGCGGACCTGTGATGGGCGCGCTCGCTGCTGCGGAGGCGC 792
Db 25906 GACCTCAGACCTACCTCGCTCCCTGCGGAGGAGGAGCTCTCATCGCC 25953

RESULT 3
US-08-456-837-6
; Sequence 6, Application US/08456837
; Patent No. 5643774
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,837
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/457,205
; FILING DATE: 01-JUN-1995
; APPLICATION NUMBER: 08/258,261
; FILING DATE: 08-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ. ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOHETICAL: NO
; ANTI-SENSE: NO
; US-08-456-837-6

Query Match 3.8%; Score 47.6; DB 1; Length 28958;
Best Local Similarity 44.1%; Pred. No. 0.007;
Matches 286; Conservative 0; Mismatches 359; Indels 3; Gaps 2;

QY 148 GGCATGAGGTGCTCATCCAGGAGGCGGAGGAGGCTCGGCTTATCCAGCAGCGGAT 207
Db 25306 GCCCAGAGACCTTACACCTCCGCGGATGTGGCTATTGCTGGCCACACCGCGGCTCAC 25365
QY 208 TTCAAGCGGCGAGCGGCAACCTGTGCGGACCGCGGAGGAGTGGGCGAGCGTGTAT 267
Db 25366 TTGAGAGACCGGCGGCTCTCGTAGCCACAAACGGAGAGAGTCTCTCGGCGCTGAC 25425
QY 268 TTATTGCTCAAGGTCAAGACCGATAGGCGGGAATAGCGCGGCTCGGAGACGAGGAG 327
Db 25426 TCGCTCGCCAGGACCAAGCGCGCGGAGACCGCTCTCGGAGGAGGAGGAGGAGGAG 25485
QY 328 ATCTGTTCAGCTTCTTGATTTGGCGGCTGACGATGCTTGCACGAGTGTGTGAT 387
Db 25486 AACTGCTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 25545
QY 388 TCGGAGCAGCAGTCAATTCCTTACGAGACGCTCAGACGCGGCGGCGGCTTACCCCTG 447
Db 25546 CTGACCTCTCGCGCGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 25605
QY 448 CTGCGCGGAGTCAAGCAGTTCGCGGCTCGACTGCGCGGCGGCGGCGGCTTACACAC--C 505
Db 25606 CAGGTGAGTGGAGCCTGCTCGCGGCTCTGCGGCGGCGGAGGAGGCGCGGCTTCTTCT 25665
QY 506 TGATGCAACCCCAAGGAGGCGGCGGCTGCTGATGGGCGGCGGCGGCGGCTTCAACCGG 565
Db 25666 CGGCTGAGCTCTTACAGCCCGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 25725
QY 566 CCGAGC-TGGTGTGATGCGGCGGCGGCGGCGGCTTACAGCAGCAGCGGCTTACCGCAAC 624
Db 25726 CGCTCGCTCGGCGGTAGAGCGCGGCGGCGGCGGCTTCTTCTTCTTCTTCTTCTTCT 25785
QY 625 GCGATGGGCGGCGGCTTACGCTTACAGATCAGATCAGATCAGATCAGATCAGATCAG 684
Db 25786 GCGCTGCTGCGAGCGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 25845
QY 685 GCGGAGTTCGCGGCGGCGGATCCACACTCGCTACTCATCGGCTTACAGAGTTCAGAGT 744
Db 25846 AAGCGCTTACACACCTTCCCGGCAAGGAGGCGGCGGCGGCTTACAGAGTTCAGAGT 25905
QY 745 GTCAACGTCGCGGACCTGTGATGGGCGGCGGCTTCTTCTTCTTCTTCTTCTTCTTCT 25953
Db 25906 GACCTCAGACCTACCTCGCTCCCTGCGGAGGAGGAGCTCTCATCGCC 25953

RESULT 4
US-08-457-342-6
; Sequence 6, Application US/08457342
; Patent No. 5652898
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25

QY 268 TTATTGCTCAAGTCTCAAGAACCGATAGCGGGGAAATACGGCCCTCGGACACGGGCGAG 327
 Db 25426 TCGCTCGCCGACGACGACCGCCCGGAGACCGCTCCCTCGGACGAGGAGCCACCGGCG 25485
 QY 328 ATCTTGTCAGCTTTGCTGATTTGGCCCGGTCAGCTGCTGACGAGATGCGTGTGGAT 387
 Db 25486 AAGCTGCTGTTGCTTTCTTCTGAGGAGAGCTGCAAGTGGGAAGGATGGCCCTCTGG 25485
 QY 388 TCCGGGACACGTCATTTGCTTACGAGACCGTCCAGACCGCGGACGAGCGACATCCCTG 447
 Db 25546 CTCGACTCTGCGCGCTTTCGCGCTCAGTCAAGATGGAGCGGCGCTGCTCCT 25605
 QY 448 CTGGCCCGGAGAGGAGTGGCGGCTGACTGCGCGCCGAGTGGCGCTTACAC--C 505
 Db 25606 CAGCTGAGTGGAGCTCTCCCTCTGCGCCGAGGAGGCGCCCTCCCTGAC 25665
 QY 506 TGATCGAACCAAGGAGGCGCGGCTGCTGATGGGCGGCTGCGGCGTCAACCGG 565
 Db 25666 CGCGTCGAGCTGTCAGACCGCCCTCTTTCGCTCATGCTTCCCTGCGGCGCTCTGG 25725
 QY 566 CCGAGC-TGCTGTGATGCGCGCGGACCGCGGCTACACGACCGCGATCGCAA 624
 Db 25726 CGCTGCTCGGCGTAGAGCCCGCGCTGCTGCGGACAGTCAAGGAGATCGCGCG 25785
 QY 625 GGCATGGGCGGACCGCTAGCTTTAGACATCAACATCGAACAACTTCGCACTCGAC 684
 Db 25786 GCTTGTGCGGAGGCGCTCTCTCTCTGAGAGCGGCGCGCGCTGCGCGCTGAGC 25845
 QY 685 GCCGAGTCTGCGGCGGATCCACACGCTACTCATCGGCGCTTACGAGTGGGTGCG 744
 Db 25846 AAGGCGCTACACCGTGGCGGCAAGGCGGCAATGGCGCGCTGAGCTGGCGCTTCC 25905
 QY 745 GTCAACGTCGCCGACCTGCTGATTGGGCGCTCTGCTGCGGCGGCG 792
 Db 25906 GACCTCGAGCTACTCTGCTCTGCGGCGACAGCTCTCATCGCC 25953

RESULT 6
 US-08-458-076A-6
 Sequence 6, Application US/08458076A
 Patent No. 5698425
 GENERAL INFORMATION:
 APPLICANT: Schupp, Thomas
 APPLICANT: Ligon, James M.
 APPLICANT: Beck, James Joseph
 APPLICANT: Hill, Dwight Steven
 APPLICANT: Ryals, John Andrew
 APPLICANT: Gaffney, Thomas Deane
 APPLICANT: Lam, Stephen Ting
 APPLICANT: Hammer, Phillip E.
 APPLICANT: Uknes, Scott Joseph
 TITLE OF INVENTION: Genes for the synthesis of
 TITLE OF INVENTION: antipathogenic substances
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Ciba-Geigy Corporation
 STREET: 7 Skyline Drive
 CITY: Haverlorne
 STATE: NY
 COUNTRY: USA
 ZIP: 10532
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/458, 076A
 FILING DATE: 01-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/457, 205

FILING DATE: 01-JUN-1995
 APPLICATION NUMBER: 08/258,261
 FILING DATE: 08-Jun-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Elmer, James Scott
 REGISTRATION NUMBER: 36,129
 REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8614
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 28958 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 ANTI-SENSE: NO
 US-08-458-076A-6

Query Match 3.84; Score 47.6; DB 2: Length 28958;
 Best Local Similarity 44.18; Pred. No. 0.0074;
 Matches 286; Conservative 0; Mismatches 359; Indels 3; Gaps 2;

QY 148 GGCATGAGTGTCTATCCAGCAGTGGCGGAGAGGCGTGGCTATACCGAGCGGAT 207
 Db 25306 GCCACAGCAGACCTCACCCTCGCGGATGTGCTATTGCTGCGACACCGCGCCAC 25365
 QY 208 TTCA GCGGCGAGCGCGCAACTGCTGGGCAACCGCGACGAGTGTGGCGAGTAT 267
 Db 25366 TTCCGACACCGCGCTCTCTGAGCCACACCGAGCGAGCTCTCTCGCGCTGAC 25425
 QY 268 TTATTGCTCAAGTCTTGCATTTGGCGGCTACGCTGCTTGCACCGATGGTGTGGAT 387
 Db 25426 TCGCTCGCCGAGCAGACCGCCCGGAGACCGCTCTCTCGGAGCGAGCGGAACCGAC 25485
 QY 328 ATCTTGTCAGCTTTGCTGATTTGGCCCGGTCAGCTGCTTGCACCGATGGTGTGGAT 387
 Db 25486 AAGCTGCTGTTGCTTTCTTCTGAGGAGCTGCAAGTGGGAAGGATGGCCCTCTGG 25545
 QY 388 TCCGGGACACGTCATTTGCTTACGAGACCGTCCAGACCGCGGCGGCTTACCGCTG 447
 Db 25546 CTCGACTCTGCGCGCTCTCTCTGAGAGCTTCAAGCATGGAGCGCGGCTGCTCT 25605
 QY 448 CTGGCCCGGAGAGGAGTGGCGGCTGACTGCGCGCCGAGTGGCGCTTACAC--C 505
 Db 25606 CAGCTGAGTGGAGCTCTCCCTCTGCGCGGAGGAGGCGCGCCCTCTGAC 25665
 QY 506 TGATCGAACCAAGGAGGCGCGGCTGCTGATGGGCGGCTGCGGCGTCAACCGG 565
 Db 25666 CGCGTCGAGCTGTCAGACCGCCCTCTTTCGCTCATGCTTCCCTGCGGCGCTCTGG 25725
 QY 566 CCGAGC-TGCTGTGATGCGCGCGGACCGCGGCTACACGAGCGCCATCGCCAC 624
 Db 25726 CGCTGCTCGGCGTAGAGCCCGCGCTGCTGCGGCGGAGTGGCGGCTTACCGCG 25785
 QY 625 GGCATGGGCGGACCGTTCAGCTTTAGACATCAACATCGAACAACTTCGCAACTGAC 684
 Db 25786 GCTTGTGCGGAGCGCTCTCTCTGAGAGACCGCGCGGCTGCGGCGCTCTGAGC 25845
 QY 685 GCCGAGTCTGCGGCGGATTCACACTGCTACTCATCGGCGCTTACAGCTGAGGCTGCG 744
 Db 25846 AAGCGCTACACCGTGGCGGCAAGGCGGCAATGGCGCGCGCTGAGCTGGCGCTCC 25905
 QY 745 GTCAACGTCGCCGACCTGCTGATTGGGCGCTCTGCTGCGGCGGCG 792
 Db 25906 GACCTCGAGCTACTCTGCTCTGCGGCGACAGCTCTCATCGCC 25953

RESULT 7
 US-08-764-233A-4


```

? Sequence 4, Application US/08764233A
? Patent No. 5716849
?
? GENERAL INFORMATION:
? APPLICANT: Ligon, James M.
? APPLICANT: Schupp, Thomas
? APPLICANT: Beck, James J.
? APPLICANT: Hill, Dwight S.
? APPLICANT: Neff, Snezana
? APPLICANT: Ryals, John A.
?
? TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen
?
? NUMBER OF SEQUENCES: 10
?
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Ciba-Geigy Corporation
? STREET: 520 White Plains Road, P.O. Box 2005
? CITY: Tarrytown
? STATE: NY
?
? COUNTRY: USA
?
? ZIP: 10591
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
?
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/764,233A
? FILING DATE:
?
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/729,214
? FILING DATE: 09-OCT-1996
?
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/258,261
? FILING DATE: 08-JUN-1994
?
? ATTORNEY/AGENT INFORMATION:
? NAME: Meigs, J. Timothy
? REGISTRATION NUMBER: 38,241
? REFERENCE/DOCKET NUMBER: 1506/CIP6
?
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (919) 541-8587
? TELEFAX: (919) 541-8689
?
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 28958 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? ORIGINAL SOURCE:
? ORGANISM: Sorangium cellulosum
? IMMEDIATE SOURCE:
? CLONE: p98/1
?
? US-08-764-233A-4

```

TELEPHONE: 919-541-8614
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 28958 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 US-08-457-335A-6

Query Match 3.8%; Score 47.6; DB 2; Length 28958;
 Best Local Similarity 44.1%; Pred. No. 0.0074;
 Matches 286; Conservative 0; Mismatches 359; Indels 3; Gaps 2;

QY 148 GCCCATGAGTGTCTCATTCACGAGAGTCCCGGAGGCGCTATCACCAGCGGAT 207
 DB 25306 GCCCAGCAGCAGCTCAGCTCCGCGGATGTGCTTATCCGCTGACACCCCGCCAC 25365
 QY 208 TTCAGGCGGAGGCGGAGTGTGCGACCGCGGAGGCGGCTGAGCGGCTGAT 267
 DB 25366 TTGAGAGCAGCGCGCGCTCTCTGAGCCACACCGGAGGCTCTCTCGCGCTGAC 25425
 QY 268 TTATGCTCAGGTCAAGAAAGCAGATAGCGCGGAAATACGCGCGCTGACAGGCGAG 327
 DB 25426 TCGCTCGCCAGAGAACCCCGCCGAGACCGTCTCTGAGGAGGAGGAGGAGGAG 25485
 QY 328 ATCTGTACGCTTCTTGATTTGGCGCGCTCAGAGCTTGCACCGATGCTTTGGAT 387
 DB 25486 AAGCTGCTTCTTCTTCTTCTGAGGAGGCTGAGGAGGAGGAGGAGGAGGAGGAG 25545
 QY 388 TCCGAGCAGCAGTCAATGCTCAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 447
 DB 25546 CTGAGCTCTCTGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 25605
 QY 448 CTGCGCCGATGAGAGGAGTCCGCGCTGAGTCCGCGCGCGCGCGCGCGCTTACAC 505
 DB 25606 CAGCTGAGAGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25665
 QY 506 TGAATGAGACCCAAAGGAGGCGGAGTGTGCTGATGAGGCGGCGCGCGCTGAGAC 565
 DB 25666 CGGCTGAGCTCTTACAGCCCGCTCTTTTCCGTATGCTCTCTGCGCGCTCTG 25725
 QY 566 CCGAGC-TGCTGTGAGTCCGCGCGCGGAGCAGCGCGCTTACAGCAGCGCGCTGAG 624
 DB 25726 CGCTCGCTGCGCTGAGAGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 25785
 QY 625 GCGATGCGCGCGAGCTTACGCTTACAGATACATGACAACTTGGGCACTGAG 684
 DB 25786 GCGTGTGCGAGGCGCTCTCTCTGAGAGGAGCGCGCGCTGCGCGCTGCGCAG 25845
 QY 685 GCCGATGCTGCGCGCGAGTACACACTGCTACTCATCGGCTGAGAGTCC-AGGCTG 744
 DB 25846 AAGGCGCTACACAGCCTGCGCGGAGAGGCGCGAGTGGCGCGCTGAGCTGCG 25905
 QY 745 GTCAACAGTGGCAGCTGTGATGAGGCGCGCTCTGCT-ACAGCGCG 792
 DB 25906 GACCTCAAGACTTCTCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 25953

RESULT 9

US-08-729-214-6
 Sequence 6, Application us/08729214
 Patent No. 5817502

GENERAL INFORMATION:

APPLICANT: Ligon, James M.
 APPLICANT: Hill, Dwight Steven
 APPLICANT: Ryals, John Andrew
 APPLICANT: Hammer, Phillip E.
 APPLICANT: van Pee, Karl-Heinz

APPLICANT: Kliner, Sabine
 TITLE OF INVENTION: Genes for the synthesis of
 TITLE OF INVENTION: antipathogenic substances
 NUMBER OF SEQUENCES: 27
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Ciba-Geigy Corporation
 STREET: 520 White Plains Road
 CITY: Tarrytown
 STATE: NY
 COUNTRY: USA
 ZIP: 10591
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/729,214
 FILING DATE: TBA
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Meigs, J. Timothy
 REGISTRATION NUMBER: 38,241
 REFERENCE/DOCKET NUMBER: CGC 1506/CIP5
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8587
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 28958 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 US-08-729-214-6

Query Match 3.8%; Score 47.5; DB 2; Length 28958;
 Best Local Similarity 44.1%; Pred. No. 0.0074;
 Matches 286; Conservative 0; Mismatches 359; Indels 3; Gaps 2;

QY 148 GCCCATGAGTGTCTCATTCACGAGAGTCCCGGAGGCGCTATCACCAGCGGAT 207
 DB 25306 GCCCAGCAGCAGCTCAGCTCCGCGGATGTGCTTATCCGCTGACACCCCGCCAC 25365
 QY 208 TTCAGGCGGAGGCGGAGTGTGCGACCGCGGAGGCTTACAGCAGCGCTGAG 267
 DB 25366 TTGAGAGCAGCGCGCGCTCTCTGAGGAGGCTGAGGAGGAGGAGGAGGAGGAG 25425
 QY 268 TTATGCTCAGGTCAAGAAAGCAGATAGCGCGGAAATACGCGCGCTGAGAGGCGAG 327
 DB 25426 TCGCTCGCCAGAGAGTCCCGCGCGGAGCAGCGCTGCTGAGAGGAGGAGGAGGAG 327
 QY 328 ATCTGTACGCTTCTTGATTTGGCGCGCTCAGAGCTTGCACCGATGCTTTGGAT 387
 DB 25486 AAGCTGCTTCTTCTTCTTCTGAGGAGGCTGAGGAGGAGGAGGAGGAGGAGGAG 25545
 QY 388 TCCGAGCAGCAGTCAATGCTCAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 447
 DB 25546 CTGAGCTCTCTGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 25605
 QY 448 CTGCGCCGATGAGAGGAGTCCGCGCTGAGTCCGCGCGCGCTTACAGCAGCGCTGAG 505
 DB 25606 CAGCTGAGAGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25665
 QY 506 TGAATGAGACCCAAAGGAGGCGGAGTGTGCTGATGAGGCGGCGCGCTGAGACCG 565
 DB 25666 CGGCTGAGCTCTTACAGCCCGCTCTTTTCCGTATGCTCTCTGCGCGCTCTG 25725
 QY 566 CCGAGC-TGCTGTGATGCGCGCGGAGCAGCGCGCTTACAGCAGCGCGCGCTGAGCAG 624

Db 25726 CGCTGCTGCGGGGTAGAGCCGCGCGCTGCTGCGCCACAGTCAGGCGGAGATCGCCGCC 25785
Qy 625 GGCATGGGCGGCGGCGTTCAGTACATCAACATGACAAACTTGGGCACTGCAC 684
Db 25786 GCGTTGCTGCGAGGCGCTCTCTCCCTCGAGAGAGGCGGCGCCGCTGCGCCAC 25845
Qy 685 GCGGAGTTCGCGGCGGATTCACACTGCTTACTCATCGGCTTACAGAGCTGAGGCTGCC 744
Db 25846 AAGGCGCTACACGCTGCGCGGCAAGGCGGCGGCGGCGGCTGAGAGCTGCGGCGCTCC 25905
Qy 745 GTCMAAGCTGCGGAGCTGATGAGGCGGCGGCTGCTGCTGCGGAGGCGCC 792
Db 25906 GACCTTCAGACACTGCTGCTGCTGCGGCGGAGAGGCTGCTGATGCGC 25953

RESULT 10
US-08-764-233A-1
Sequence 1, Application US/08764233A
Patent No. 5716849
GENERAL INFORMATION:
APPLICANT: Ligon, James M.
APPLICANT: Schupp, Thomas
APPLICANT: Beck, James J.
APPLICANT: Hill, Dwight S.
APPLICANT: Neff, Suzanna
APPLICANT: Ryals, John A.
TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/764,233A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/729,214
FILING DATE: 09-06-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: 1506/CIP6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 49377 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Sorangium cellulosum
IMMEDIATE SOURCE:
CLONE: p98/1, pUL3, and pVKM15
FEATURE:
NAME/KEY: misc.feature
LOCATION: 383..760
OTHER INFORMATION: /product= "SorB"
OTHER INFORMATION: /note= "This gene encodes a protein that is highly homologous

OTHER INFORMATION: the reductase domains of type I PKs such as eryA from
OTHER INFORMATION: Saccharopolyspora erythraea."
FEATURE:
NAME/KEY: misc.feature
LOCATION: 927..19874
OTHER INFORMATION: /product= "Sora"
OTHER INFORMATION: /note= "Gene product is highly homologous to type I PKs t
OTHER INFORMATION: are known to be involved in the synthesis of polyketide
OTHER INFORMATION: compounds."
FEATURE:
NAME/KEY: misc.feature
LOCATION: 942..7115
OTHER INFORMATION: /product= "Module 1 of Sora"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 7203..12884
OTHER INFORMATION: /product= "Module 2 of Sora"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 13455..19616
OTHER INFORMATION: /product= "Module 3 of Sora"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 19871..46318
OTHER INFORMATION: /product= "SorB"
OTHER INFORMATION: /note= "Gene product is highly homologous to type I PKs ge
FEATURE:
NAME/KEY: misc.feature
LOCATION: 19870..24556
OTHER INFORMATION: /product= "Module 1 of SorB"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 24638..30820
OTHER INFORMATION: /product= "Module 2 of SorB"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 30881..35446
OTHER INFORMATION: /product= "Module 3 of SorB"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 35528..40114
OTHER INFORMATION: /product= "Module 4 of SorB"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 40190..46318
OTHER INFORMATION: /product= "Module 5 of SorB"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 46851..47891
OTHER INFORMATION: /product= "SorB"
OTHER INFORMATION: /note= "The protein encoded by the sorB gene is highly
OTHER INFORMATION: homologous to the methyltransferase from Streptomyces
OTHER INFORMATION: hygroscopicus that is involved in the synthesis of the
OTHER INFORMATION: polyketide rapamycin."
US-08-764-233A-1

Query Match 3 84; Score 47.6; DB: 2; Length 49377;
Best Local Similarity 44.18; Pred. No. 0.0086;
Matches 286; Conservative 0; Mismatches 359; Indels 3; Gaps 2;

Qy 148 GGCATGAGTGTCTATCATGAGGAGTGCAGAGAGGCTGATACCGAGCGGAT 207
Db 41609 GCCACGACGACCTACCTGCGGAGATGTGCTATTCGCTGCGACACCGCGCCAC 41668
Qy 208 TTCAAGGCGGAGCGGCAACTGTGCGACCGCGGAGGAGTGTGCGGAGCGGTAT 267
Db 41669 TTCAGGACCGCGGCGCTCTGTAGCCACACGCGAGAGCTCTCTCGGCTGAC 41728
Qy 268 TTATGCTCAAGTCAAGAAACCATAGCGGCGGAATACGCGCGCTGCGACAGGGCAG 327
Db 41729 TCGCTGCGGCGGAGAGCGCGCGCGGAGAGCAGCTCTCTGAGAGGAGGAGAACCGAGCG 41788

QY 338 ATCTGTCACGCTTCTTGCAATTTGGCCGCTCAGTGTCTGACAGCGATGCTTGTGAT 387
 Db 41789 AAGCTGCTTGTGTTCTTCTGCGAGAGGCTGCAATGGGATGGCCCTCTGCTG 41848
 QY 388 TCCGACACGCTCAATTTGCTTCCAGAGACCGTCCAGACCGCCGACGCGCTACCTG 447
 Db 41849 CTCGACTCTCCCGCTCTCCGCGCTCAGCTCGAAGCTGAGGCGCCGCTGCTCT 41908
 QY 448 CTTGGCCCGATGAGGAGTGGCCGCTGACAGCGCCGCTGAGCGCTTACACAC--C 505
 Db 41909 CACGTGAGTGAAGCTGCTGCGCTGCTGCGCCGCGAGAGAGGCGCCCTCTCTGAC 41968
 QY 506 TGAATGGAACCCAGAGGCGCGGCTGTGCTGATGAGGCGGCTGCGCGCTGAGACCG 565
 Db 41969 CGCGTGAAGCTGATAGAGCGCCCTCTTGGCGTATGTTGCTCTCCCTGGCGGCTCTG 42028
 QY 566 CCGAGC-TGCTGTGTCGCGCGCGCGCGCGCGCTACAGCAGCGCGCTGCGGAC 624
 Db 42029 CCGTGTGCTGCGGCTGAGAGCGCGCGCGCGCTGCGCGCAGTCAAGGCGCGCGCG 42088
 QY 625 GGCATGCGCGCGCGCTGAGGCTTCTAGATCAATCAATGAGCAACTTGGGCACTGAC 684
 Db 42089 GCGTGTGCTGCGCGCGCTGCTCTCTGAGAGCGCGCGCGCTGCGCGCTGCGGAC 42148
 QY 685 GCGGATGCTGCGCGCGCTGCTCTCTGAGAGCGCGCGCGCTGCGCGCTGCGGAC 744
 Db 42149 AAGGCGCTGCGCGCGCTGCGCGCGCGCGCGCGCTGCGCGCTGCGGCGCTGCG 42208
 QY 745 GTCGAACGTCGCGCGCTGCTGATGAGGCGCGCTGCTGCTGCGCGCGCGCG 792
 Db 42209 GACTCTCAGACCTGCTGCTGCTGCGCGCGCGCGCGCGCTGCTGCTGCGCG 42256

RESULT 11

US-08-816-105A-2
 ; Sequence 2, Application US/08816105A

; Patent No. 589882

; GENERAL INFORMATION:

; APPLICANT: Crueger, Anneliese; Dellweg, Hans-Georg; Lenz,

; APPLICANT: Jurgen; Schroder, Werner; Pape, Hermann;

; APPLICANT: Gecke, Klaus; Schaper, Beate; Hemker, Michael;

; APPLICANT: Piepersberg, Wolfgang; Distler, Jurgen;

; APPLICANT: Stralman, Ansgar

; TITLE OF INVENTION: PROCESSES FOR PREPARING ACARBIOSYL

; TITLE OF INVENTION: TRANSFERASE AND FOR USING IT IN THE

; TITLE OF INVENTION: CONVERSION OF ACARBOSE HOMOLOGUES

; TITLE OF INVENTION: INTO ACARBOSE, FOR THE PREPARATION

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sprung Kramer Schaefer & Briscoe

; STREET: 660 White Plains Road

; CITY: Tarrytown

; STATE: New York

; COUNTRY: USA

; ZIP: 10591-5144

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage

; OPERATING SYSTEM: System 7.5

; SOFTWARE: Wordperfect 3.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/816,105A

; FILING DATE: 14-MAR-1997

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: DE 19625269.5

; FILING DATE: 25-JUN-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: DE 19611252.4

; FILING DATE: 22-MAR-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Kuitl G. Briscoe

; REGISTRATION NUMBER: 33,141
 ; REFERENCE/DOCKET NUMBER: Bayer 9814-KGB
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (914) 332-1700
 ; TELEFAX: (914) 332-1844
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2582 nucleotides
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; US-08-816-105A-2

Query Match 3.7%; Score 46; DB 4; Length 2582;
 Best Local Similarity 47.3%; Pred. No. 0.0088;
 Matches 139; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

QY 353 CCGCGTACGCTCTGACAGCATGCTTGTGATTCGCGACAGTCAATTCCTACG 412
 Db 1576 CCGCGTACGCTCTGACAGCATGCTTGTGATTCGCGACAGTCAATTCCTACG 1635
 QY 413 AAGCGTCCAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 472
 Db 1636 TGACGCTGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1695
 QY 473 GTCGACTGCG 532
 Db 1696 GCGAGTGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1755
 QY 533 TCTGATGCG 592
 Db 1756 TCGGCG 1815
 QY 593 CCGCGCGCTCAACAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 646
 Db 1816 CCGCGCGAGACCGTCCGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1869

RESULT 12

US-08-403-852D-3
 ; Sequence 3, Application US/08403852D

; Patent No. 5891695

; GENERAL INFORMATION:

; APPLICANT: Blanc, Veronique

; APPLICANT: Blanchet, Francis

; APPLICANT: Crozet, Joel

; APPLICANT: Jacques, Nathalie

; APPLICANT: Lacroix, Patricia

; APPLICANT: Thibaut, Denis

; APPLICANT: Zagorec, Monique

; APPLICANT: Debussche, Laurent

; APPLICANT: De Crey-Lagard, Valerie

; TITLE OF INVENTION: Polypeptides Involved In The

; TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences

; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Flimegan, Henderson, Farabow, Garrett & Dunner

; STREET: 1300 I Street, N.W., Suite 700

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20005-3315

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/403,852D

; FILING DATE: 10-MAY-1995

; PRIOR APPLICATION DATA:

APPLICANT: Crouzet, Joel
APPLICANT: Jacques, Nathalie
APPLICANT: Lacroix, Patricia
APPLICANT: Thibaut, Denis
APPLICANT: Zagorec, Monique
APPLICANT: Debussche, Laurent
APPLICANT: De Crey-Lagard, Valerie
TITLE OF INVENTION: Polypeptides Involved In The
TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences
TITLE OF INVENTION: Coding For These Polypeptides And Their Use
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/403,852D
FILING DATE: 10-MAY-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/FR 93/00923
FILING DATE: 25-SEP-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: FR 92/11441
FILING DATE: 25-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03806, 0054-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1833 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: S.pristinaespiralis
FEATURE:
NAME/KEY: CDS
LOCATION: 103..1689
US-08-403-852D-6

Query Match 3.4%; Score 42.4; DB 3; Length 1833;
Best Local Similarity 50.5%; Pred. No. 0.062;
Matches 103; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
DB 413 TGAATGCTGGAGCTGCTTCTGGCGCGGCTCCCTTGGGCGCTTGCACGAGACT 472
QY 533 TGGTATGGGGGGGGTGGCGGCTGGAACCGGCGACGCTGCTGGTATGGGGCGGCA 592
DB 413 TGAATGCTGGAGCTGCTTCTGGCGCGGCTCCCTTGGGCGCTTGCACGAGACT 472
QY 593 CGCGCGGTACAGACGAGCGCGCATGCGCAAGGCGATGGGCGGAGCGGCTTCTAG 652
DB 473 CGCGCGGTACAGACGAGCGCGCGCGCGCGCGCTGGGCGGCGCTGCTGCGACCA 592
QY 653 ACATCAACATCGACAACTTGGCAACTGACCGCGAGTTCGCGGCGGATCCACACTC 712
DB 533 CGACCGTGGCGGATCATCATCAGAGTCTTGCAGACGACGAGAGCGCCCGCGGATGGCC 592
QY 713 GCTACTCATGCGGCTTACGAGCTCG 736

DB 593 TGGGAGCGCGGCGACACTCATGCG 616
RESULT 15
US-08-173-508-3
Sequence 3, Application US/08173508
Patent No. 5616485
GENERAL INFORMATION:
APPLICANT: Bartfeld, Daniel
APPLICANT: Butler, Michael J.
APPLICANT: Hadary, Dany
APPLICANT: Jenish, David
APPLICANT: Krieger, Timothy
TITLE OF INVENTION: STREPTOMYCIN PROTEASES AND IMPROVED
TITLE OF INVENTION: STREPTOMYCIN STRAINS FOR EXPRESSION OF PEPTIDES AND
TITLE OF INVENTION: POLYPEPTIDES
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/173,508
FILING DATE: 23-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 18740/125/CACO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2185 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 531..2069
FEATURE:
NAME/KEY: sig-peptide
LOCATION: 531..902
FEATURE:
NAME/KEY: mat-peptide
LOCATION: 903..2069
FEATURE:
NAME/KEY: misc-feature
LOCATION: 531..533
OTHER INFORMATION: /note="Met at position -124
represents fmet"
US-08-173-508-3

Query Match 3.4%; Score 42.4; DB 1; Length 2185;
Best Local Similarity 47.4%; Pred. No. 0.066;
Matches 127; Conservative 0; Mismatches 141; Indels 0; Gaps 0;
DB 525 CGCGGTGTGCTGATGAGGGGGGCGGCTGCAACCGGCGAGCGTGTGATGG 584
DB 1250 CGCGGTGTGCTGATGAGGGGGGCGGCTGCAACCGGCGAGCGTGTGATGG 1309
QY 585 CGCGGTGTGCTGATGAGGGGGGCGGCTGCAACCGGCGAGCGTGTGATGG 644

Db 1310 CCCCAGCGGCTTCTTCTACACCGAGGCCGTGCTCTGCGCTTCATGTGGGCGCGAGCA 1369
QY 645 GGTTCATGACATCAACATCGACAACCTTCGCAACTCGACCGAGTTCTGCGCCGAT 704
Db 1370 CGGGGTGAGCTGACCAACAAGCTATTACCGACCCGTGGTACTTCAACTGCAAGGA 1429
QY 705 CCACACTCGCTACTCATCGGCTACGAGCTCGAGGGTGCCCTCAACGCTGCCGACTGGT 764
Db 1430 CGACCCCGACCAAGGCGCTCGTCGAGGCGCTCTCGGGGCTCCCGGTACGCGAGAA 1489
QY 765 GATTGGGCGCGCTCCTGCTGCGAGCGCC 792
Db 1490 GAAGGGCGCGTCAACCTCGCGCGGCC 1517

Search completed: June 22, 2000, 15:06:11
Job time: 17849 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 12:07:22 ; Search time 5541.94 Seconds
(Without alignments)
904.708 Million cell updates/sec

Title: US-09-362-485-4
Perfect score: 1237
Sequence: 1 ATCTGCGATTTATCGAAC.....CGAGCGATGATCGGCCG 1237

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 segs, 2026611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:*

1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: em_est20:*
21: em_est21:*
22: em_est22:*
23: em_est23:*
24: em_est24:*
25: em_est25:*
26: em_est26:*
27: em_est27:*
28: em_est28:*
29: em_est29:*
30: em_est30:*
31: em_est31:*
32: em_est32:*
33: em_est33:*
34: em_est34:*
35: em_est35:*
36: em_est36:*
37: em_est37:*
38: em_est38:*
39: em_est39:*
40: em_est40:*
41: em_est41:*
42: em_est42:*
43: em_est43:*
44: em_est44:*
45: em_est45:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_est48:*
83: gb_est49:*
84: gb_est50:*
85: gb_est51:*
86: em_est35:*
87: em_est36:*
88: em_est37:*
89: em_est38:*
90: gb_est52:*
91: gb_est53:*
92: gb_est54:*
93: gb_est55:*
94: gb_est56:*
95: em_est39:*
96: em_est40:*
97: em_est41:*
98: em_est42:*
99: em_est43:*
100: em_est44:*
101: em_est45:*
102: gb_est57:*
103: gb_est58:*
104: em_est46:*
105: gb_est59:*
106: gb_est60:*
107: gb_est61:*
108: gb_est62:*
109: gb_est63:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES

FEATURES					
source		Location/Qualifiers			
		1..846			
		/organism="Drosophila melanogaster"			
		/plasmid="PbeloBAC11"			
		/db_xref="taxon:7227"			
		/clone_1lb="DrosBAC"			
		/clone="BACN04N13"			
		/note="end : 77"			
BASE COUNT		300 a	47 c	40 g	175 t 284 others
ORIGIN					
Query Match 3.8%; Score 46.6; DB 82; Length 846;					
Best Local Similarity 13.0%; Pied No. 6.6;					
Matches 33; conservative 127; Mismatches 93; Indels 0; Gaps 0					
Oy	363	TGCTTCACCGAGCGCTTGTGGATCCGGACCAACGTAATTGCCTACGAACCGTGCA	422		
Db	842	TTCSTSSSTSSTCSSTSBTTSTTSTSTSASSSSSTSSSTSSSTSSSTSBTBT	783		
Oy	423	GACCGCGCAGCGCGCACACTACCTGCTGCCCGCATGACGAAAGTCGGCGGCGACTGCC	482		
Db	762	SBSSTTSTTSSSTSTSTSSSTCTSCSSSSSTTSLSASASSTSTSTTGSSBS	723		
Oy	483	CGCCACGATGTGGCGCTTACCACCTGATGCGAACCAAGGGGCGCGGTGCTGATGG	542		
Db	722	SBSTGSTSTSBTBTBTBRSBBSRSSSSSTSTPTBRSBSSSBSTSTSTSS	663		
Oy	543	CGGGGCGCGCGCGCGAACCAGCGACGTCGTCGATCGCGCGCGCGCGCGCTA	602		
Db	662	SSSSTSTSSSSSTCTTSTTSSSCSSSSSTSTSSBSSSSCRTSSCSSSSSCT	603		
Oy	603	CACGCGAGCCGC 615			
Db	602	CGCCSCGCCCCC 590			
RESULT 7					
CNS0181E/c					
LOCUS		CNS0181E 1101 bp DNA GSS 26-JUL-1999			
DEFINITION		Drosophila melanogaster genome survey sequence SP6 end of BAC			
		BACN37H05 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.			
ACCESSION		AL108764			
VERSION		AL108764.1 GI:5629068			
KEYWORDS		GSS.			
SOURCE		fruit fly.			
ORGANISM		Drosophila melanogaster			
		Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
		Mycetozoa; Neoptera; Endopterygota; Diptera; Brachycera;			
		Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
REFERENCE		1 (bases 1 to 1101)			
AUTHORS		Genoscope.			
TITLE		Direct Submission			
JOURNAL		Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - FRANCE (E-mail : segre@genoscope.cns.fr Web : www.genoscope.cns.fr)			
COMMENT		Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanoagaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.			
FEATURES		Location/Qualifiers			
source		1..1101			
		/organism="Drosophila melanogaster"			
		/plasmid="PbeloBAC11"			
		/db_xref="taxon:7227"			
		/clone_1lb="DrosBAC"			
		/clone="BACN37H05"			

BASE COUNT	61 a	215 c	301 g	235 t	289 others
ORIGIN					
Query Match	3.8%: Score 46.6; DB 83; Length 1101;				
Best Local Similarity	40.2%: Pred. No. 6.9;				
Matches 101; Conservative	40; Mismatches 109; Indels 1; Gaps 1				
QY	389	CGGACACACGTCATTTGCTCTACGAGACCGTTCACAGACCCCGCCAGCGGACGACATACCCCTCC	448		
DB	308	CCCCCCTTTTGGMAAASCCCAVASSCCSCCCSCCCSCCCGCGCGSCGACSCSCSCG	249		
QY	449	TTGCCCCGATGAGCGAAG:CGCGGTTCGACTGCGCGCCCGCAGGTGTGGCGTTACCACTGA	508		
DB	248	AAGMGSMGVEGACACAGAGGCGGAGACCGCGSCACSCCCSCSCSCSCSCSSSS	189		
QY	509	TGCGAACCCAGAGGGGCGCGGTGTGCTGATGAGGCGGGGTGCCCGGCTGCAACGGCGCG	568		
DB	188	CGCGACCCSSAGGGGCGCCSGCGGCCACCCCGCGGACGACGCGCCAC-CHMYGGGSG	130		
QY	569	AAGTGTGATGTCGCGCGCGCGCGCGCGGTTCACAGACCGCGCGCAACGCGCAACGGA	628		
DB	129	ACGCAAGGAGGAGAGGCGCGGAGCGCCGCGCCGCGCCGCGCGCGCGCGCGCGGA	70		
QY	629	TGGCGCGGACG 639			
DB	69	AGCGGCCAACG 59			
RESULT	8				
LOCUS	CNS006XK	935 bp	DNA	GSS	03-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC # BACR14N09 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AF066051				
VERSION	AF066051.1 GI:4945019				
KEYWORDS	GSS.				
SOURCE	fruit fly.				
ORGANISM	Drosophila melanogaster				
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
AUTHORS	1 (bases 1 to 935)				
TITLE	Genoscope.				
JOURNAL	Direct Submission				
COMMENT	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
FEATURES	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .				
SOURCE	Location/Qualifiers				
	1..935				
	/organism="Drosophila melanogaster"				
	/db_xref="taxon:7227"				
	/clone_lib="RPCI-98"				
	/clone="BACR14N09"				
	/note="end : T7"				
BASE COUNT	257 a	170 c	162 g	96 t	250 others

BASE COUNT	61 a	215 c	301 g	235 t	289 others
ORIGIN					
Query Match	3.8%: Score 46.6; DB 83; Length 1101;				
Best Local Similarity	40.2%: Pred. No. 6.9;				
Matches 101; Conservative	40; Mismatches 109; Indels 1; Gaps 1				
QY	389	CGGACACACGTCATTTGCTCTACGAGACCGTTCACAGACCCGCCACGCGGACGACATACCCCTTC	448		
DB	308	CCCCCCTTTTGGMAAASCCCAVASSSCCCSCCCSCCCSCCCGCGCGSCGACSCSCSCG	249		
QY	449	TTGCCCCGATGAGCAAG:GCCGGTTCAGCTGCGCGCCAGGTGTGGCGTTACCACTTA	508		
DB	248	AAGMGSMGVEGACACAGGCGCGGAGACCGCCSCCACCSCCCSCCSCCSCCSCSSS	189		
QY	509	TGCGAACCCAGAGGGGCGCGGTGTGCTGATGGCGGGGTGCCCGGTGCAACGGCGCG	568		
DB	188	CGCGACCCSSAGGGGCGCCSGCGGCCACCCCGCGGACGACGCGCAC-CCMYGGGSG	130		
QY	569	AAGTGTGATGTCGCGCGCGCGCGCGGTTCACAGACCGCCGCAACGCGCAACGGA	628		
DB	129	ACGCAAGGGGAGAGGGCGCGNAGCCCGCCGNCACCCCGCGCGCACGCGCGGGA	70		
QY	629	TGGCGCGGAC 639			
DB	69	AGCGCCCAAC 59			
RESULT	8				
CNS006XK					
LOCUS	CNS006XK	935 bp	DNA	GSS	03-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC # BACR14N09 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AU066051				
VERSION	AU066051.1 GI:4945019				
KEYWORDS	GSS.				
SOURCE	fruit fly.				
ORGANISM	Drosophila melanogaster				
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
AUTHORS	1 (bases 1 to 935)				
TITLE	Genoscope.				
JOURNAL	Direct Submission				
COMMENT	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
FEATURES	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .				
SOURCE	Location/Qualifiers				
	1..935				
	/organism="Drosophila melanogaster"				
	/db_xref="taxon:7227"				
	/clone_lib="RPCI-98"				
	/clone="BACR14N09"				
	/note="end : T7"				
BASE COUNT	257 a	170 c	162 g	96 t	250 others


```
/clone_lib="605 - Endosperm cDNA library from Schmidt lab
/tissue_type="nucellar, embryo, and endosperm"
/dev_stage="10-14 days post-pollination"
/lab_host="DH5(alpha)"
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 14:52:44 ; Search time 8627.09 Seconds

(without alignments)
-138.469 Million cell updates/sec

Title: US-09-362-485-5
1228
Sequence: 1 ATCTTGAGATTAATCGAAC.....GGAGTAAGGAGAGATGA 1228

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

GenDbml: *
1: gb_ba1: *
2: gb_ba2: *
3: gb_cm: *
4: gb_ov: *
5: gb_pat: *
6: gb_ph: *
7: gb_pl1: *
8: gb_pl2: *
9: gb_pl3: *
10: gb_pr1: *
11: gb_pr2: *
12: gb_pr3: *
13: gb_ro: *
14: gb_sts: *
15: gb_sy: *
16: gb_un: *
17: gb_vl: *
18: em_fun: *
19: em_hum1: *
20: em_hum2: *
21: em_in: *
22: em_om: *
23: em_or: *
24: em_ov: *
25: em_pat: *
26: em_ph: *
27: em_pl: *
28: em_ro: *
29: em_sts: *
30: em_sy: *
31: em_un: *
32: em_vl: *
33: gb_htg1: *
34: gb_htg2: *
35: gb_in1: *
36: gb_in2: *
37: em_ba1: *
38: em_ba2: *
39: em_hum3: *
40: gb_pr4: *
41: gb_htg3: *
42: gb_htg4: *
43: gb_htg5: *
44: gb_htg6: *

45: gb_htg7: *
46: em_htg1: *
47: em_htg2: *
48: em_htg3: *
49: em_hum5: *
50: gb_pl3: *
51: gb_pr5: *
52: gb_htg8: *
53: gb_htg9: *
54: gb_htg10: *
55: gb_htg11: *
56: gb_htg12: *
57: gb_htg13: *
58: gb_htg14: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1228	100.0	1228	5	A87608 Sequence 5
2	1228	100.0	1328	5	A89748 Sequence 5
3	1228	100.0	1237	5	A87607 Sequence 4
4	1228	100.0	1237	5	A89747 Sequence 4
5	1228	100.0	56414	1	MTV002
6	1226.4	99.9	1236	5	A87613 Sequence 10
7	1226.4	99.9	1236	5	A89753 Sequence 10
8	1224.4	99.7	1235	5	A87606 Sequence 3
9	1224.4	99.7	1235	5	A89746 Sequence 3
10	1213.4	98.8	1235	5	A87609 Sequence 6
11	1213.4	98.8	1235	5	A89749 Sequence 6
12	1212.8	98.8	1235	5	A87611 Sequence 8
13	1212.8	98.8	1235	5	A89751 Sequence 8
14	1211.8	98.7	1229	5	A87610 Sequence 7
15	1211.8	98.7	1229	5	A89750 Sequence 7
16	1210.4	98.6	1245	5	A87605 Sequence 2
17	1210.4	98.6	1245	5	A89745 Sequence 2
18	1210.4	98.6	1260	5	A87604 Sequence 1
19	1210.4	98.6	1260	5	A89744 Sequence 1
20	1210.4	98.6	2584	1	MTALADH
21	1206.4	98.2	1209	5	A89752 Sequence 9
22	1199.4	97.7	1208	5	A87612 Sequence 9
23	1192.4	97.1	1194	2	MT092472
24	698	56.8	720	5	A89759 Sequence 16
25	538.4	43.8	40745	1	SC151
26	467.6	38.1	682	5	A89758 Sequence 15
27	467.6	38.1	682	5	A89752 Sequence 19
28	377	30.7	8046	2	AF049107 Myxococ
29	351.2	28.6	11514	2	AE002028 Deinococ
30	331.4	27.0	2780	1	AE238118 Rhizobium
31	330.8	26.9	1585	2	AF070715 Vibrio pr
32	312.8	25.5	1736	1	M33299 B. steartohe
33	293	23.9	1477	2	AF070715 Shewanell
34	288.4	23.5	2113	1	D37807 Phormidium
35	277.2	22.6	1125	5	E01847 DNA encodin
36	258	21.0	137740	1	D90900 Synchocyst
37	243.2	19.8	1375	1	AB013821 Enterobac
38	243	19.8	1776	1	BACALDH
39	235.6	19.2	1709	1	M33298 B. sphaericu
40	235.6	19.2	12316	1	L20916 Bacillus su
41	235.6	19.2	217420	1	B5282015 B. subtilis
42	218.2	17.8	1714	2	Z99120 Bacillus su
43	107.2	8.7	10632	2	AF070714 Carnobact
44	99.2	8.1	10512	2	AE001565 Helicobac
45	74.2	6.0	3160	1	AE000639 Helicobac
					RU05294 Rhodospirill

ALIGNMENTS

RESULT 1
 LOCUS A87608 1228 bp DNA PAT 22-JAN-2000
 DEFINITION Sequence 5 from Patent WO9836089.
 ACCESSION A87608
 VERSION A87608.1 GI:6736248
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified
 REFERENCE 1 (bases 1 to 1228)
 AUTHORS Flohe, L. and Singh, M.
 TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
 JOURNAL Patent: WO 9836089-A 20-AUG-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source 1..1228
 location/Qualifiers
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 236 a 391 c 382 g 219 t
 ORIGIN

Query Match 100.0%; Score 1228; DB 5; Length 1228;
 Best Local Similarity 100.0%; Pred. No. 4.2e-169;
 Matches 1228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATCTTGCAATTAATCGAATCTTCTTCACTAGAGCGTACAGTATCGAGAGGGGTATC 60
 1 ATCTTGCAATTAATCGAATCTTCTTCACTAGAGCGTACAGTATCGAGAGGGGTATC 60
 61 ATGGCGCTCGGTATCCGACCGGAGACCAAAACACGAATTCGGGCTGGCCATCACCCG 120
 61 ATGGCGCTCGGTATCCGACCGGAGACCAAAACACGAATTCGGGCTGGCCATCACCCG 120
 121 GCCGCGCTCGGCAACTACCCGCTGCGCATGAGTGCATCCAGGAGGTGCGGCA 180
 121 GCCGCGCTCGGCAACTACCCGCTGCGCATGAGTGCATCCAGGAGGTGCGGCA 180
 121 GCCGCGCTCGGCAACTACCCGCTGCGCATGAGTGCATCCAGGAGGTGCGGCA 180
 121 GCCGCGCTCGGCAACTACCCGCTGCGCATGAGTGCATCCAGGAGGTGCGGCA 180
 181 GAGGCGCTCGGTATCCGACCGGAGGATTCAGGCGGCGGCAACTGCTGGCACC 240
 181 GAGGCGCTCGGTATCCGACCGGAGGATTCAGGCGGCGGCAACTGCTGGCACC 240
 181 GAGGCGCTCGGTATCCGACCGGAGGATTCAGGCGGCGGCAACTGCTGGCACC 240
 241 GCCGACCAAGGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 300
 241 GCCGACCAAGGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 300
 301 GAATACGCGCGCTGCGACACGGGCGGAGATCTTTCACGTTCTTGCATTTGGCGCGTCA 360
 301 GAATACGCGCGCTGCGACACGGGCGGAGATCTTTCACGTTCTTGCATTTGGCGCGTCA 360
 361 CGGCTTGCAACCGATGCGGTGTTGGATTCGGGCGGCGGCGGCGGCGGCGGCGGCGG 420
 361 CGGCTTGCAACCGATGCGGTGTTGGATTCGGGCGGCGGCGGCGGCGGCGGCGGCGG 420
 421 CAGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
 421 CAGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
 481 GCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
 481 GCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
 541 GGGGCGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
 541 GGGGCGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
 601 TACAAAGCAGCCCGCATCGGCAACGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
 601 TACAAAGCAGCCCGCATCGGCAACGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
 661 ATGCAAAAGTGTGGGCAAGTGAAGCGGCGGAGTTTGGGCGGCGGAGTCCACACTGCTACTCA 720

661 ATGCAAAAGTGTGGGCAAGTGAAGCGGCGGAGTTTGGGCGGCGGAGTCCACACTGCTACTCA 720
 721 TCGGCTACGAGCTCGAGGGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
 721 TCGGCTACGAGCTCGAGGGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
 781 GTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
 781 GTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
 841 GTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
 841 GTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
 901 ACCACCTAGACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
 901 ACCACCTAGACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
 961 ATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
 961 ATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
 1021 GTGCTGAGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
 1021 GTGCTGAGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
 1081 AAAGGCTTTGCGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
 1081 AAAGGCTTTGCGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
 1141 GTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
 1141 GTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
 1201 CACACGTCGGGAGTAAGCAGGATGA 1228
 1201 CACACGTCGCTAGTAAGGAGAGCGATGA 1228

RESULT 2
 LOCUS A89748 1228 bp DNA PAT 22-JAN-2000
 DEFINITION Sequence 5 from Patent WO9832862.
 ACCESSION A89748
 VERSION A89748.1 GI:6738292
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified
 REFERENCE 1 (bases 1 to 1228)
 AUTHORS Flohe, L. and Singh, M.
 TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
 JOURNAL Patent: WO 9832862-A 30-JUL-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source 1..1228
 location/Qualifiers
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 236 a 391 c 382 g 219 t
 ORIGIN

Query Match 100.0%; Score 1228; DB 5; Length 1228;
 Best Local Similarity 100.0%; Pred. No. 4.2e-169;
 Matches 1228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATCTTGCAATTAATCGAATCTTCTTCACTAGAGCGTACAGTATCGAGAGGGGTATC 60
 1 ATCTTGCAATTAATCGAATCTTCTTCACTAGAGCGTACAGTATCGAGAGGGGTATC 60
 61 ATGGCGCTCGGTATCCGACCGGAGACCAAAACACGAATTCGGGCTGGCCATCACCCG 120

No.	Score	Match	Length	DB	ID	Description
1	57.4	4.6	925	82	CNS0091P	AL05013 Drosophila
2	55.2	4.5	925	82	CNS0091P	AL05013 Drosophila
3	51.8	4.2	535	74	AV390505	AV390505 Drosophila
4	49.9	4.0	910	82	CNS0060N	AL056652 Drosophila
5	46.6	3.8	844	82	CNS0052P	AL056652 Drosophila
6	46.6	3.8	846	82	CNS010RJ	AL099337 Drosophila
7	46.6	3.8	1101	83	CNS0181E	AL108764 Drosophila
8	46.4	3.8	935	82	CNS006X	AL066051 Drosophila
9	45.8	3.7	935	82	CNS006X	AL066051 Drosophila
10	45.4	3.7	591	65	AM128786	AM128786 Drosophila
11	45.4	3.7	744	83	CNS0172K	AL108460 Drosophila
12	45.2	3.7	1101	83	CNS0175X	AL108460 Drosophila
13	45	3.6	315	51	AV1712257	AV1712257 Drosophila
14	45	3.6	846	82	CNS010RJ	AL099337 Drosophila
15	44.8	3.6	932	82	CNS00720	AL066742 Drosophila
16	44	3.6	843	82	CNS00C51	AL059666 Drosophila
17	43.8	3.5	427	51	AT736678	AT736678 Drosophila
18	43.8	3.5	512	74	AV396794	AV396794 Drosophila
19	43.8	3.5	521	74	AV397116	AV397116 Drosophila
20	43.8	3.5	526	74	AV394827	AV394827 Drosophila
21	43.8	3.5	645	82	CNS01213	AL101589 Drosophila
22	43.6	3.5	1101	83	CNS016DE	AL107216 Drosophila
23	43.2	3.5	439	91	AO864063	AO864063 Drosophila
24	43.2	3.5	798	82	CNS000AJA	AO864063 Drosophila
25	43	3.5	413	74	AV396947	AV396947 Drosophila
26	41.8	3.4	414	80	AM286084	AM286084 Drosophila
27	41.8	3.4	465	80	AM287007	AM287007 Drosophila
28	41.8	3.4	530	74	AV396353	AV396353 Drosophila
29	41.8	3.4	839	82	CNS004NB	AL054280 Drosophila
30	41.8	3.4	932	82	CNS00720	AL066742 Drosophila
31	41.8	3.4	1101	83	CNS0175X	AL108460 Drosophila
32	41.6	3.4	506	74	AM180713	AM180713 Drosophila
33	41.6	3.4	870	82	CNS006E2	AL064271 Drosophila
34	41.4	3.4	1101	82	CNS00LXJ	AL078875 Drosophila
35	41.4	3.4	1201	83	CNS014BJ	AL103945 Drosophila
36	41.2	3.3	639	60	AO1812194	AO1812194 Drosophila
37	41.2	3.3	774	105	AO3272321	AO3272321 Drosophila
38	41.2	3.3	910	82	CNS0060N	AL056652 Drosophila
39	41	3.3	605	46	AI461529	AI461529 Drosophila
40	40.8	3.3	384	64	AM054773	AM054773 Drosophila
41	40.8	3.3	385	21	D48746	D48746 Drosophila
42	40.8	3.3	437	74	AV395790	AV395790 Drosophila
43	40.6	3.3	506	74	AV396405	AV396405 Drosophila
44	40.4	3.3	574	91	AO864186	AO864186 Drosophila
45	40.2	3.3	916	83	CNS0135Q	AL105860 Drosophila

ALIGNMENTS

RESULT 1
LOCUS CNS0091P 925 bp DNA 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TE73 end of BAC #
BACR19D16 of RPI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL053013.1 GI:4934461
VERSION AL053013.1
KEYWORDS fruit fly,
SOURCE Drosophila melanogaster
ORGANISM Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 925)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage;
BP 191 91006 Evry cedex - FRANCE (E-mail: seqref@genoscope.cns.fr
COMMENT - Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a

collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazuto Osoegawa and
Aaron Mammeter in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
location/Qualifiers
1..925
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPI-98"
/clone="BACR19D16"
/note="end : TE73"

BASE COUNT 120 a 61 c 61 g 172 t 511 others
ORIGIN

Query Match 4.6%; Score 57.4; DB 82; Length 925;
Best Local Similarity 11.9%; Pred. No. 0.06;
Matches 43; Conservative 176; Mismatches 142; Indels 0; Gaps 0;

QY 441 CCCCTGCTTCCCGATGACGACGCGGTCGACGCGCCGAGTGGCCCTTAC 500
DB 920 SCSCSCSBCSSSMTSSTSSBSCSSSSTSSMSBSSSSSSSSSGTS 861
QY 501 CACCTGATGCGAACGAGCGCGGCGGTCGATGCGGCGGTCGCCGCGA 560
DB 860 SACKKMASSCCCGCGACBCCSSSSCCGSASARVKNASGACRGGSGASA 801
QY 561 CCGGCGACGCTGCTGATGCGCGCGCGCGCGGCTTACGACGCGCGATGCC 620
DB 800 SSSSSACBSSSSSSSCASGWSASSSSSSSSSSSSSSSSSSSSSSSS 741
QY 621 AAGGCGATGGCGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 680
DB 740 VSSASSSSSSSSVSCSVSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 681
QY 681 GACCGCGATGCTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 740
DB 680 SCSTASMSAARSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSG 621
QY 741 GCCGTCAAGCTCCGACCTGATGATGGCGCGGCGGCGGCGGCGGCGGCG 800
DB 620 GSGSVASSGMSVSSGSGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 561
QY 801 A 801
DB 560 R 560

RESULT 2
LOCUS CNS0091P 925 bp DNA 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TE73 end of BAC #
BACR19D16 of RPI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL053013.1 GI:4934461
VERSION AL053013.1
KEYWORDS fruit fly,
SOURCE Drosophila melanogaster
ORGANISM Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 925)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage;
BP 191 91006 Evry cedex - FRANCE (E-mail: seqref@genoscope.cns.fr
COMMENT - Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a

AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seque@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Oseguwa and Atron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
Source
1..925
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACR19D16"
/note="End : TET3"

BASE COUNT 120 a 61 c 61 g 172 t 511 others

ORIGIN

Query Match 4.5%; Score 55.2; DB 82; Length 925;
Best Local Similarity 11.8%; Pred. No. 0.16;
Matches 41; Conservative 170; Mismatches 137; Indels 0; Gaps 0;

424 CCGCGCAGCGCCATACACCTGCTTGCCCGCATGAGCGCAATGCGCGGTGACATCGCG 483
Db 577 SCGCGSSSCCBCCGCCSCSSYCCSSBSSSSKCSSTBSGCCSCSSKSVCTGCGSSSS 636
484 CCCAGGTGGCCCTTACCAACCTGATGCGAACCAAGGGGGCCGCGTGTGATGGGCG 543
Db 637 CASSSSSTSSSSTSSSTSSKSSSSSSSSSYTTKSTKSAAGSGSWSAGGGSGTGS 696
QY 544 GGGTCCCGCGGCTCGAACCGCCGACATGCTGTATGATCGGCCCGGACCGCGGTACA 603
Db 697 SSSSSSSTSSSVSSVSGKSTBSBSGSSBSGSSSSSTSSBSCTSTSSSSSSSS 756
QY 604 ACGGAGCCCGCATGCGCAAGCGATGGGGGCGACCGTTACGTTTACAGATCAACATCG 663
Db 757 TCCCTCCCSYSSSTSSSTSSSTSMGTSBSGSSSVGTSSSDSTSTCCSCCYMCTCS 816
QY 664 ACAAACTTCGCGACATCGACGCGGATCTGTGGGCGCGGATCCACATCGCTACTATCG 723
Db 817 TYBMBCTYSTSGGSSSSGKGKGYKCGGGCGSSSTNMBTSSACSSSSSSSSSVS 876
QY 724 CCTACGAGCTGAGGGTGCCGTCAAACGTGCGGACCTGGTATTGGG 771
Db 877 SSSKSSASBSVSSSGSGSVSSNSASSSSSSGSVSSGSGSGSGSVS 924

RESULT 3
AV390505 AV390505 535 bp mRNA EST 21-JAN-2000
LOCUS AV390505 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
DEFINITION cDNA clone CM052B05_r, mRNA sequence.
ACCESSION AV390505
VERSION AV390505.1 GI:6544721
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales; Chlamydomonas.
1 (bases 1 to 535)

AUTHORS Asamizu, E., Nakamura, Y., Sato, S., Fukuzawa, H. and Tabata, S.
 TITLE A Large Scale Structural Analysis of cDNAs in a Unicellular Green
 Alga, Chlamydomonas reinhardtii. I. Generation of 3433
 Non-redundant Expressed Sequence Tags
 JOURNAL DNA Res. 6, 369-373 (1999)
 COMMENT On Dec 20, 1995 this sequence version replaced gi:11355919.
 Contact: Yasukazu Nakamura
 The First Laboratory for I. Lu. Gene Research
 Kazusa DNA Research Institute
 Yata 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/
 FEATURES
 SOURCE
 1. 535
 /organism="Chlamydomonas reinhardtii"
 /strain="C9"
 /db_xref="taxon:3055"
 /clone="CM052B05_r"
 /clone_11b="Chlamydomonas reinhardtii C9"
 /dev_stage="photoautotrophic growth"
 /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
 XhoI"
 BASE COUNT 107 a 195 c 148 g 85 t
 ORIGIN
 Query Match 4.2%; Score 51.8; DB 74; Length 535;
 Best Local Similarity 51.5%; Pred. No. 0.66;
 Matches 119; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
 Oy 564 GCCGACGTCTGTGTATCGCGCCGGCACCCGGGCTTCACAGCGACCCGGCATCGCCAC 623
 Db 221 GCTGACGAGGTGGCCATCGACAGCTGCTCTCTCCATGACGAGGCTGACCGCGGTG 280
 Oy 624 GGCATGGCGCGCACCGCTTACGCTTACGATCAATGACATGACAATTCGGGACCTGCAC 683
 Db 281 CGCCGTGGGTGACCGCTGTGCTCCCAAGGCCCTTCAGTAGACCAAGCGGCGGCTGCAC 340
 Oy 684 GCCGAGTTCTG:GCGCGGATCCACACTGCTACTCATCGGCGCTTACGAGCTGAGGCTGCC 743
 Db 341 GCGGACTGGAAACCGCGCTCA1 ACCACTGCTTTCACAGCGGCGCTCCACATGCTGCAGGGC 400
 Oy 744 GTCAACGTGCCGACCTGTGATTTGGGCGCGCTCTGTGTGCCAGGGGCCACAG 794
 Db 401 GCGGCTCTGGCGCTGCTGCAGCGGAGGGGCTCAAGATGACGCGCATG 451
 RESULT 4
 CINS0060N/C
 LOCUS CINS0060N 910 bp DNA GSS 03-JUN-1999
 DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC #
 BACR14J21 of RPCL-98 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 ACCESSION AL065629
 VERSION AL065629.1 GI:4944698
 KEYWORDS GSS.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE
 TITLE Direct Submission
 AUTHORS Submitted (03-JUN-1999) Genoscope - Centre National de Sequencage : :
 JOURNAL BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 COMMENT
 - Web : www.genoscope.cns.fr)
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see <http://www.fruitfly.org> The BDGP Drosophila
 melanogaster BAC library was prepared by Kazuhiro Osoegawa and
 Atron Mammoser in Pieter de Jong's laboratory in the Department of

FEATURES	pBelobAC11.									
SOURCE	Location/Qualifiers									
	1. 846									
	/organism="Drosophila melanogaster"									
	/plasmid="pBelobAC11"									
	/db_xref="taxon:7227"									
	/clone_11b="DrosBAC"									
	/clone="BACN04N13"									
	/note="end : 17"									
BASE COUNT	300	a	47	c	40	g	175	t	284	others
ORIGIN										
Query Match	3.8%; Score 46.6; DB 82; Length 846;									
Best Local Similarity	13.0%; Pred. No. 7.1;									
Matches	33;	Conservative	127;	Mismatches	193;	Indels	0;	Gaps	0;	
OY	362	TCCTTCACCGATGCGTGTGTGGATTCCGGACACACGTCATATGCTTACGAGACCGTCCA	421							
Db	842	TTCTCTSSSTSSSTSCSSSSSBTBTSTTTSTSTSTSTSTSTSTSTSTSTSTSTSTSTBTBT	783							
OY	422	GACCCGCGACGCGACGACGCTGCTGTCGCCCGACGACGCGAGTGGCGGCGACTGCG	481							
Db	782	SBSTTTTSTTSSSTTSTTSSSTCTCTSSCSCSSSTTSSSASASTSTSTTGTCSBS	723							
OY	482	CGCCACGATGCGCGTTACCACTGATGCAACCCAGGCGCGCGGTGCTGATGG	541							
Db	722	SBSTGTSSTSBTBTBTBTSSBSBSSSSSSTBTBTSSBSBSSTSTSTSTSTST	663							
OY	542	CGGGGTGCGCGCGCTGCAACCGCGACGCTGTGTGATCGCGCGCGACCGCGGCTTA	601							
Db	662	SSSSBTSTSTSSBSSTCTTSTTSSSCSSSSSTSTSSBSBSBSSTSTSSSSSCT	603							
OY	602	CAACGACGCGCGC	614							
Db	602	CGCCCGSCCGCC	590							
RESULT	7									
LOCUS	CNS0181E/c									
DEFINITION	Drosophila melanogaster genome survey sequence, SP6 end of BAC BACN37H05 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.									
ACCESSION	AL108764									
VERSION	AL108764.1 GI:5629068									
KEYWORDS	GSS.									
SOURCE	fruit fly.									
ORGANISM	Drosophila melanogaster									
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.									
REFERENCE	1 (bases 1 to 1101)									
AUTHORS	Genoscope.									
TITLE	Direct Submission									
JOURNAL	Submitted (23-JUN-1999) Genoscope - Centre National de Sequençage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequefgenoscope.cns.fr									
COMMENT	- Web : www.genoscope.cns.fr) - Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (BDGP) - http://www.egdp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CNRS (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.									
FEATURES	Location/Qualifiers									
SOURCE	1. 1101									
	/organism="Drosophila melanogaster"									
	/plasmid="pBelobAC11"									
	/db_xref="taxon:7227"									
	/clone_11b="DrosBAC"									
	/clone="BACN37H05"									

BASE COUNT		61 a	215 c	301 g	235 t	289 others
ORIGIN						
Query Match		3.8%	Score 46.6;	DB 83;	Length 1101;	
Best Local Similarity		40.2%;	Pred. No. 7.3;			
Matches 101; Conservative		40;	Mismatches 109;	Indels 1;	Gaps 1;	
OY	388	CCGGACCAAGCTGATTCATTCCGTACACAGACCCTGCACACCGCCGACGGCGCATCACCCTTGC	447			
Db	308	CCCCCCTVMMVMVGMMAAACCCCCCNAAVCSSCCSCSSCCSSCCSSCCSACSACSSCCSG	249			
OY	448	TTCGCCCATGATGAGGAMAATCCCGCTGCTACTCGCCGACCGCCAGTTGGCCCTTACCACTGA	507			
Db	248	AAGMGSMGVGGACAAGGGCGGAGACCCGCGCACCCSCSCACCCSSSCGCCGqscscsssss	189			
OY	508	TGCGAACCACAAAGGGCGCCGCGTGCTGATGAGGGGGGGTGGCCGCGTCGAACCGCGCG	567			
Db	188	CGCACACCCSSAGGGGCGCCSCCCGCCACCCCGCCGSACVACCGCCAC-CMVGGSg	130			
OY	568	ACGTGTGTGATCGCGCCCGCACCGCCGCGGTACACAGACCGCCGATTCGCAACGCA	627			
Db	129	ACCGAAGAAGGAGAGAGGGCGGAGNAGCCCGCNCACACCCCGCGCCGCGCCACGCGCGGCA	70			
OY	628	TGGCGCGCACC	538			
Db	69	AGCGCCCAACC	59			
RESULT	8					
LOCUS	CNS006XK	935 bp	DNA	GSS	03-JUN-1999	
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC #					
ACCESSION	BACR14N09 of RPCI-98 library from Drosophila melanogaster (fruit					
VERSION	AL066051.1	GI:4945019				
KEYWORDS	GSS.					
SOURCE	fruit fly.					
ORGANISM	Drosophila melanogaster					
REFERENCE	Eukaryota; Meta.:Arthropoda; Tracheata; Hexapoda; Insecta;					
AUTHORS	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;					
TITLE	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.					
JOURNAL	1 (bases 1 to 935)					
COMMENT	Genoscope. Direct Submission Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr - Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm. Location/Qualifiers 1..935 /organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone_1id="RPCI-98" /clone="BACR14N09" /note="end : T7"					
FEATURES	source					
BASE COUNT		257 a	170 c	162 g	96 t	250 others

/db_xref="taxon:7955"
/clone_lib="zebrafish Mashu M1MG EST"
/sex="mixed"
/tissue_type="26 somite embryos, adult livers, shield
stage embryos"
/lab_host="XLI-blue MRF"
/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; 1st
strand cDNA was primed with a Not I - oligo(dT)15 primer
(5'-GACATGATTCAGTACGAGCGGCGCCCTTTTCTTTT3');
double-stranded cDNA was ligated into Sal I adaptors (BRU),
digested with Not I and cloned into the Not I and Sal I
sites of the pSPORT1 vector (BRU). Library was constructed
by Matthew Clark (Lehrach lab; ICRF, London and Max Planck
Institut fuer Molekulare Genetik, Berlin). cDNAs for EST
analysis were selected following oligonucleotide
hybridization fingerprinting of arrayed clones from
zebrafish late somitogenesis (26 ss), adult liver or
embryonic shield stage (5.6 h) libraries. Fingerprint
data were used to computationally cluster cDNAs, and a
single cDNA from each cluster was chosen for sequencing.
In some cases multiple members of the same cluster were
sequenced to assess clustering parameters or single clones
were sequenced additional times to assess quality
control."

BASE COUNT 119 a 169 c 156 g 147 t
ORIGIN

Query Match 3.7%; Score 45.4; DB 69; Length 591;
Best Local Similarity 50.2%; Pred. No. 11;
Matches 112; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

OY 2 TCTTCAGATTAATGAACTTCTTCACATGAGCGTATATGAGAGGGGTATCA 61
DB 185 TCTTAGAGCATATCAAGCCCTCAACCGCTCAGTATGAGCATCTCTTCAAAACAC 244
OY 62 TCCGCGTGGTATTCGACGAGACCAAAACACGATTCGGGTGGCCATCACCCCG 121
DB 245 TCACCGTCGCGCTCTTCAAGAAATCTTCAAGATGAGCGTGGAGGCGCATCTCTCTG 304
OY 122 CCGGGTGGCGGACATCAACCGCTCGTGGCATGAGTGTCTATCCAGGAGGTGGCGAG 181
DB 305 CCGGAGTGAACCTCTCATCAAGACGAGGCTTAAATGTGGTGGAGTCCGGCGGAG 364
OY 182 AGGCGTCCGCTATCACGCGGAGATTTCAAGCGGACGCGC 224
DB 365 ATCTGCCAAGTCTCTGATGATGATGATACCAAGCAGGAGC 407

RESULT 11
LOCUS CNS0172K 744 bp DNA GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence Sp6 end of BAC
BACN37B10 of Drosophila library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL108698
VERSION AL108698.1 GI:5629002
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 744)
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seq@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre

d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelosBAC11.

FEATURES
source Location/Qualifiers
1..744
/organism="Drosophila melanogaster"
/plasmid="pBelosBAC11"
/db_xref="taxon:7227"
/clone_lib="DrosBAC"
/clone="BACN37B10"
/note="end : Sp6"

BASE COUNT 20 a 191 c 245 g 105 t 183 others
ORIGIN

Query Match 3.7%; Score 45.4; DB 83; Length 744;
Best Local Similarity 33.8%; Pred. No. 12;
Matches 92; Conservative 42; Mismatches 138; Indels 0; Gaps 0;

OY 312 CCTGGACAGCGGCGCATCTGTTACGTTCTTGCAATTTGGCGGTCAGTCTGCAC 371
DB 175 CCCTGAGATCTTCCTCTTCCTTGTGCTGCTGCGGCGGCTGCGCTGCTGCTG 234
OY 372 GATGCGTGTGATTCGCGACACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 431
DB 235 SYGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 294
OY 432 GCGGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 491
DB 295 SCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 354
OY 492 GCGGCTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 551
DB 355 NNCCNCCG 414
OY 552 GCGCTGCAACCG 583
DB 415 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 446

RESULT 12
LOCUS CNS0175Y 1101 bp DNA GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence Sp6 end of BAC
BACN37L08 of Drosophila library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL108460
VERSION AL108460.1 GI:5628764
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seq@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelosBAC11.

COMMENT
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/plasmid="pBelosBAC11"

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 14:54:29 ; Search time 8627.09 Seconds
(without alignments)
-138.582 Million cell updates/sec

Title: US-09-362-485-7
Perfect score: 1229
Sequence: 1 ATCTTGCAATATATGAC.....GAGTAGGAGAGCATGATG 1229

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

GenBank: *
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl3:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_to:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vl:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl1:*
27: em_pl2:*
28: em_pl3:*
29: em_sy:*
30: em_un:*
31: em_vl:*
32: gb_hg1:*
33: gb_hg2:*
34: gb_hg3:*
35: gb_hg4:*
36: gb_hg5:*
37: em_ba1:*
38: em_ba2:*
39: em_hum3:*
40: gb_pr4:*
41: gb_hg3:*
42: gb_hg4:*
43: gb_hg5:*
44: gb_hg6:*

45: gb_hg7:*
46: em_hg1:*
47: em_hg2:*
48: em_hg3:*
49: em_hg4:*
50: gb_pl3:*
51: gb_pr5:*
52: gb_hg8:*
53: gb_hg9:*
54: gb_hg10:*
55: gb_hg11:*
56: gb_hg12:*
57: gb_hg13:*
58: gb_hg14:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1228	99.9	1229	5	A87610 Sequence 7
2	1228	99.9	1229	5	A89750 Sequence 7
3	1228	99.9	1235	5	A87611 Sequence 8
4	1228	99.9	1235	5	A89751 Sequence 8
5	1235.4	99.7	1235	5	A87609 Sequence 6
6	1235.4	99.7	1235	5	A89749 Sequence 6
7	1215.4	98.9	1236	5	A87613 Sequence 10
8	1215.4	98.9	1236	5	A89753 Sequence 10
9	1213.8	98.8	1237	5	A87607 Sequence 4
10	1213.8	98.8	1237	5	A89747 Sequence 4
11	1213.8	98.8	56414	1	MTV002
12	1213.4	98.7	1235	5	A87606 Sequence 3
13	1213.4	98.7	1235	5	A89746 Sequence 3
14	1211.8	98.6	1228	5	A87608 Sequence 5
15	1211.8	98.6	1228	5	A89748 Sequence 5
16	1199.4	97.6	1245	5	A87605 Sequence 2
17	1199.4	97.6	1245	5	A89745 Sequence 2
18	1199.4	97.6	1260	5	A87604 Sequence 1
19	1199.4	97.6	1260	5	A89744 Sequence 1
20	1199.4	97.6	2584	1	MTA1ADH
21	1193.4	97.1	1209	5	A87612 Sequence 9
22	1187.4	96.6	1208	5	A89752 Sequence 9
23	1180.4	96.0	1194	2	MTU92472
24	684.4	55.7	720	5	A89759 Sequence 16
25	524.8	42.7	40745	1	SC151
26	454	36.9	682	5	A89758 Sequence 15
27	454	36.9	682	5	A89752 Sequence 15
28	363.4	29.6	8046	2	AF049107 Myxococcus
29	339.2	27.6	11514	2	AE002028 Deinococcus
30	318.8	25.9	1585	2	AF070716 Vibrio pr
31	317.8	25.9	2780	1	RLE238118 Rhizobium
32	300.8	24.5	1736	1	BACALDHC
33	279.4	22.7	1477	1	AF070715 Shewanella
34	275.4	22.5	2113	1	D37807 Phormidium
35	265.2	21.6	1125	5	E01847
36	249.2	20.3	137740	1	D90900 Synecocyst
37	232.8	18.9	1375	1	AB013821 Enterobac
38	231	18.8	1776	1	BACALDH
39	223.6	18.2	1709	1	BACALDH
40	223.6	18.2	12316	1	B6282015 B.subtilis
41	223.6	18.2	217420	1	BSUB0017 Bacillus su
42	206.2	16.8	1714	2	AF070714 Carnobact
43	101.6	8.3	10652	2	AE001565 Helicobac
44	93.6	7.6	10512	2	AE000639 Helicobac
45	74	6.0	3160	1	RR005294 Rhodospirillum

ALIGNMENTS

```

RESULT 1
LOCUS AB7610 1229 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 7 from Patent WO9836089.
ACCESSION AB7610
VERSION AB7610.1 GI:6736250
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1229)
AUTHORS Flohe, L. and Singh, M.
TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL Patent: WO 9836089-A-20-AUG-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
Source 1..1229
/db_xref="taxon:32644"
BASE COUNT 236 a 391 c 382 g 219 t 1 others
ORIGIN
Query Match 99.9%; Score 1228; DB 5; Length 1229;
Best Local Similarity 100.0%; Pred. No. 2.7e-167;
Matches 1229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTTGAGATTATCGAATCTTCTTACACTGAGCGTACAGTATGAGAGGGGTATC 60
DB 1 ATCTTGAGATTATCGAATCTTCTTACACTGAGCGTACAGTATGAGAGGGGTATC 60
QY 61 ATGCGCGTGGGATTCGACCGAGACCAAAAACAAGATTCCGGGTGCCATACCCCG 120
DB 61 ATGCGCGTGGGATTCGACCGAGACCAAAAACAAGATTCCGGGTGCCATACCCCG 120
QY 121 GCGCGCGTGGGAGACTAACCGGTGTGGCATGAGTGTCTATCCAGGAGTCCGGA 180
DB 121 GCGCGCGTGGGAGACTAACCGGTGTGGCATGAGTGTCTATCCAGGAGTCCGGA 180
QY 181 GAGGCGTGGGATTCAGCGAGCGGATTTCAAGGGCGGCGCACTGGTGGCGAC 240
DB 181 GAGGCGTGGGATTCAGCGAGCGGATTTCAAGGGCGGCGCACTGGTGGCGAC 240
QY 241 GCCGACAGGTGTGGCGGAGCGTATTATTGCTCAAGGTCAAGAACCGATAGCGCG 300
DB 241 GCCGACAGGTGTGGCGGAGCGTATTATTGCTCAAGGTCAAGAACCGATAGCGCG 300
QY 301 GAATACGCGCGCTGCGACAGGGGAGTCTGTACGTTCTTGATTTGGCCGCTAC 360
DB 301 GAATACGCGCGCTGCGACAGGGGAGTCTGTACGTTCTTGATTTGGCCGCTAC 360
QY 361 GTGCTTGACCGATCGATTGTGATTCGGGACCACTCAATTGCTACAGAGACGCTCC 420
DB 361 GTGCTTGACCGATCGATTGTGATTCGGGACCACTCAATTGCTACAGAGACGCTCC 420
QY 421 AGACCGCGAAGGCGCACTACCCCTGTGCGCCGATGAGGAAAGTCCCGTGCATCG 480
DB 421 AGACCGCGAAGGCGCACTACCCCTGTGCGCCGATGAGGAAAGTCCCGTGCATCG 480
QY 481 CCGCGCGATTTGGCGCTTACCACTGATGCGAACCCAGGGGCGCGGCTGTGCTGATGG 540
DB 481 CCGCGCGATTTGGCGCTTACCACTGATGCGAACCCAGGGGCGCGGCTGTGCTGATGG 540
QY 541 GCGGGGTGCGCGGCTGCAACCGGCGAGTGTGATGGGCGCGGACCGCGGCT 600
DB 541 GCGGGGTGCGCGGCTGCAACCGGCGAGTGTGATGGGCGCGGACCGCGGCT 600
QY 601 ACAACGAGCCCGCATCGCAACGGGAGTGGGCGGCGAGTACGGTTCTAGACATCAACA 660
DB 601 ACAACGAGCCCGCATCGCAACGGGAGTGGGCGGCGAGTACGGTTCTAGACATCAACA 660
QY 661 TCGACAAACTTCGGAATCGACGCGGAGTTCTGGGCGCGGATCCACACTGCTACTCAT 720

```

```

DB 661 TCGACAAACTTCGGAATCGACGCGGAGTTCTGGGCGCGGATCCACACTGCTACTCAT 720
QY 721 CCGGCTACGAGCTCGAGGGTCCGCTCAACGTGCCGACTGTGATTTGGGCGCTCTGG 780
DB 721 CCGGCTACGAGCTCGAGGGTCCGCTCAACGTGCCGACTGTGATTTGGGCGCTCTGG 780
QY 781 TGCCAGCGCGCAAGGCGCAATTAATCTGGAATTCATTTGCGGCAATGAAACCGAG 840
DB 781 TGCCAGCGCGCAAGGCGCAATTAATCTGGAATTCATTTGCGGCAATGAAACCGAG 840
QY 841 GTGGGCTACTGTGATATAGCCATCGACGAGGCGGCTGTTTGAAGGCTCAGACCGA 900
DB 841 GTGGGCTACTGTGATATAGCCATCGACGAGGCGGCTGTTTGAAGGCTCAGACCGA 900
QY 901 CCACCTACGACACCGGAGCTTGGCGGTGCACGACACGCTGTTTACTGGTGGCAACA 960
DB 901 CCACCTACGACACCGGAGCTTGGCGGTGCACGACACGCTGTTTACTGGTGGCAACA 960
QY 961 TGCCGCGCTGGGTCCGAAAGCTGACCTACGGGCTGACCAAGCGAGATGCGGTATG 1020
DB 961 TGCCGCGCTGGGTCCGAAAGCTGACCTACGGGCTGACCAAGCGAGATGCGGTATG 1020
QY 1021 TGCTGAGCTTCCCGACCATGCTGGCGGGCGGTGCGCGGTGAAATCCGCACTAGCA 1080
DB 1021 TGCTGAGCTTCCCGACCATGCTGGCGGGCGGTGCGCGGTGAAATCCGCACTAGCA 1080
QY 1081 AAGGTCTTTGAGCGACGACGAGGGGCTTACTGTCCGACGAGGTGGCCACCGACTGGGG 1140
DB 1081 AAGGTCTTTGAGCGACGAGGGGCTTACTGTCCGACGAGGTGGCCACCGACTGGGG 1140
QY 1141 TGCCGCTACCGAGCCCGGAGCGTGTGCTGACCTCTGGCGGCTCGTTACGCGGAG 1200
DB 1141 TGCCGCTACCGAGCCCGGAGCGTGTGCTGACCTCTGGCGGCTCGTTACGCGGAG 1200
QY 1201 ACACGTCNGAGTAAAGGAGCATGATG 1229
DB 1201 ACACGTCNGAGTAAAGGAGCATGATG 1229

RESULT 2
LOCUS AB9750 1229 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 7 from Patent WO9832862.
ACCESSION AB9750
VERSION AB9750.1 GI:6738284
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1229)
AUTHORS Flohe, L. and Singh, M.
TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
JOURNAL Patent: WO 9832862-A-30-JUL-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
Source 1..1229
/db_xref="taxon:32644"
BASE COUNT 236 a 391 c 382 g 219 t 1 others
ORIGIN
Query Match 99.9%; Score 1228; DB 5; Length 1229;
Best Local Similarity 100.0%; Pred. No. 2.7e-167;
Matches 1229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTTGAGATTATCGAATCTTCTTACACTGAGCGTACAGTATGAGAGGGGTATC 60
DB 1 ATCTTGAGATTATCGAATCTTCTTACACTGAGCGTACAGTATGAGAGGGGTATC 60
QY 61 ATGCGCGTGGGATTCGACCGAGACCAAAAACAAGATTCCGGGTGCCATACCCCG 120

```



```

Oy 601 ACAACGACCCCGCATGCGCAACGGCATGGGCGGACCGCTTACGGTCTAGACATCAACA 660
    |||||||
Db 601 ACAACGACCCCGCATGCGCAACGGCATGGGCGGACCGCTTACGGTCTAGACATCAACA 660
Oy 661 TCGACAAACTTGGGCACTCGACGGCGGATTCGCGGCGGATCGACACTCGCTACTAT 720
    |||||||
Db 661 TCGACAAACTTGGGCACTCGACGGCGGATTCGCGGCGGATCGACACTCGCTACTAT 720
Oy 721 CGGCGTACGAGCTCGAGGCTGGCGCTCAACGTCGCGGCTGATGATGGGCGGCTCTGG 780
    |||||||
Db 721 CGGCGTACGAGCTCGAGGCTGGCGCTCAACGTCGCGGCTGATGATGGGCGGCTCTGG 780
Oy 781 TCGCAGGCGCCAAAGCACCACCAATTAGTCTCGAATTCATCTCGCGCATATGAACCAAG 840
    |||||||
Db 781 TCGCAGGCGCCAAAGCACCACCAATTAGTCTCGAATTCATCTCGCGCATATGAACCAAG 840
Oy 841 GTGGGCTACTGCTGATATAGCCATCGACAGGGGCGCTGTTTGAAGGCTCAGACCGA 900
    |||||||
Db 841 GTGGGCTACTGCTGATATAGCCATCGACAGGGGCGCTGTTTGAAGGCTCAGACCGA 900
Oy 901 CCACCTACGACCCACCGGAGCTGCGCGTGCAGACAGCGCTGTTTACTGCGTGAAGCA 960
    |||||||
Db 901 CCACCTACGACCCACCGGAGCTGCGCGTGCAGACAGCGCTGTTTACTGCGTGAAGCA 960
Oy 961 TCGCCGCGCTGCGTGGCGGAAGAGCTGACCTACGCGCTGACCAACGCGAGATCGCGTATG 1020
    |||||||
Db 961 TCGCCGCGCTGCGTGGCGGAAGAGCTGACCTACGCGCTGACCAACGCGAGATCGCGTATG 1020
Oy 1021 TCGCTGAGCTTCCGACCATGCTGCGGGGCGGCTGCGCGTGCMAATCGGACACTAGCCA 1080
    |||||||
Db 1021 TCGCTGAGCTTCCGACCATGCTGCGGGGCGGCTGCGCGTGCMAATCGGACACTAGCCA 1080
Oy 1081 AAGGCTTTGACGACGACGAAGGGGCGTACTGTCGAAACGGGTGGCGACCGACTGGGG 1140
    |||||||
Db 1081 AAGGCTTTGACGACGACGAAGGGGCGTACTGTCGAAACGGGTGGCGACCGACTGGGG 1140
Oy 1141 TCGCGTTCACGAGCGCGCGACGCTGCTGCGCTGACTCTCGGCGCTGTTACGGCGGAC 1200
    |||||||
Db 1141 TCGCGTTCACGAGCGCGCGACGCTGCTGCGCTGACTCTCGGCGCTGTTACGGCGGAC 1200
Oy 1201 ACACGTGNGAGTAGGGAAGGATGATG 1229
    |||||||
Db 1201 ACACGTGNGAGTAGGGAAGGATGATG 1229

RESULT 4
LOCUS A89751 1235 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 8 from Patent WO9832862.
ACCESSION A89751
VERSION A89751.1 GI:6738285
KEYWORDS
SOURCE .
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1235)
AUTHORS Flohe, L. and Singh, M.
TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
JOURNML Patent: WO 9832862-A 30-JUL-1998:
FEATURES
    FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
    Location/Qualifiers
        1..1235
            /organism="unidentified"
            /db_xref="taxon:32644"
BASE COUNT 236 a 394 c 385 g 220 t
ORIGIN
    Query Match 99.9%; Score 1228; DB 5; Length 1235;
    Best Local Similarity 99.9%; Pred. No. 2,7e-167;
    Matches 1228; Conservative 0; Mismatches 1; Indels 0; Gaps 0:
Oy 1 ATCTTGACGATTAATGCAACTTTCTTACACTGAAAGCGTACGATGAGAGGGGTAATC 60

```

```

    |||||||
Db 1 ATCTTGACGATTAATGCAACTTTCTTACACTGAAAGCTACGATGAGAGGGGTAATC 60
Oy 61 ATGCGCGTGGTATTCGACCGAGACCAAAAACAAGAAATTCGCGG;GGCCATACCCCG 120
    |||||||
Db 61 ATGCGCGTGGTATTCGACCGAGACCAAAAACAAGAAATTCGCGG;GGCCATACCCCG 120
Oy 121 GCGGCGTGGGGAATACCCGCTGCGCATATAGTGCATTCAGGCAAGAGTGGCCGGA 180
    |||||||
Db 121 GCGGCGTGGGGAATACCCGCTGCGCATATAGTGCATTCAGGCAAGAGTGGCCGGA 180
Oy 181 GAGGCGTGGCTATCACCGACGCGGATTTCAAGGCGGAGCGGCAAACTGGTGGCACC 240
    |||||||
Db 181 GAGGCGTGGCTATCACCGACGCGGATTTCAAGGCGGAGCGGCAAACTGGTGGCACC 240
Oy 241 GCGGACAGTGTGGGCGGACGCTGATTTATGCTCAAGGTCAAAGAACGATAGCGCG 300
    |||||||
Db 241 GCGGACAGTGTGGGCGGACGCTGATTTATGCTCAAGGTCAAAGAACGATAGCGCG 300
Oy 301 GAATAGGCGCGCTGGCGACGGGCGATCTTGTACAGTCTTGCAATTTGGCGCGCTAC 360
    |||||||
Db 301 GAATAGGCGCGCTGGCGACGGGCGATCTTGTACAGTCTTGCAATTTGGCGCGCTAC 360
Oy 361 GTGCTTGACCGATGCTGTTGGATTCCGGCACACGTCMAATTCCTTACGAGACGCTCC 420
    |||||||
Db 361 GTGCTTGACCGATGCTGTTGGATTCCGGCACACGTCMAATTCCTTACGAGACGCTCC 420
Oy 421 AGACCGCGGAAGGGCGCACTACCCCTGCTGCGCGGATGAGCAAGTGGCGGTGACTCG 480
    |||||||
Db 421 AGACCGCGGAAGGGCGCACTACCCCTGCTGCGCGGATGAGCAAGTGGCGGTGACTCG 480
Oy 481 CCGCCAGAGTTGGGCGCTTACCACTGATGGCAACCAAGGGGCGGCGGTGGTGTATGG 540
    |||||||
Db 481 CCGCCAGAGTTGGGCGCTTACCACTGATGGCAACCAAGGGGCGGCGGTGGTGTATGG 540
Oy 541 GCGGCGTGGCGGCGTGGCAACCGGCGACGCTGCTGATGAGCGCGCGCACCGCGGCT 600
    |||||||
Db 541 GCGGCGTGGCGGCGTGGCAACCGGCGACGCTGCTGATGAGCGCGCGCACCGCGGCT 600
Oy 601 ACAACGACCCCGCATGCGCAACGGCATGGGCGGACCGCTTACGCTTACATCAACA 660
    |||||||
Db 601 ACAACGACCCCGCATGCGCAACGGCATGGGCGGACCGCTTACGCTTACATCAACA 660
Oy 661 TCGACAAACTTGGGCACTCGACGGCGGATTCGCGGCGGATCGACACTCGCTACTAT 720
    |||||||
Db 661 TCGACAAACTTGGGCACTCGACGGCGGATTCGCGGCGGATCGACACTCGCTACTAT 720
Oy 721 CGGCGTACGAGCTCGAGGCTGGCGCTCAACGTCGCGGCTGATGATGGGCGGCTCTGG 780
    |||||||
Db 721 CGGCGTACGAGCTCGAGGCTGGCGCTCAACGTCGCGGCTGATGATGGGCGGCTCTGG 780
Oy 781 TCGCAGGCGCCAAAGCACCACCAATTAGTCTCGAATTCATCTCGCGCATATGAACCAAG 840
    |||||||
Db 781 TCGCAGGCGCCAAAGCACCACCAATTAGTCTCGAATTCATCTCGCGCATATGAACCAAG 840
Oy 841 GTGGGCTACTGCTGATATAGCCATCGACAGGGGCGCTGTTTGAAGGCTCAGACCGA 900
    |||||||
Db 841 GTGGGCTACTGCTGATATAGCCATCGACAGGGGCGCTGTTTGAAGGCTCAGACCGA 900
Oy 901 CCACCTACGACCCACCGGAGCTGCGCGTGCAGACAGCGCTGTTTACTGCGTGAAGCA 960
    |||||||
Db 901 CCACCTACGACCCACCGGAGCTGCGCGTGCAGACAGCGCTGTTTACTGCGTGAAGCA 960
Oy 961 TCGCCGCGCTGCGTGGCGGAAGAGCTGACCTACGCGCTGACCAACGCGAGATCGCGTATG 1020
    |||||||
Db 961 TCGCCGCGCTGCGTGGCGGAAGAGCTGACCTACGCGCTGACCAACGCGAGATCGCGTATG 1020
Oy 1021 TCGCTGAGCTTCCGACCATGCTGCGGGGCGGCTGCGCGTGCMAATCGGACACTAGCCA 1080
    |||||||
Db 1021 TCGCTGAGCTTCCGACCATGCTGCGGGGCGGCTGCGCGTGCMAATCGGACACTAGCCA 1080
Oy 1081 AAGGCTTTGACGACGACGAAGGGGCGTACTGTCGAAACGGGTGGCGACCGACTGGGG 1140
    |||||||

```

Db 1081 AAGTCTTTCGACGACGACGAGGCGCTACTGTCCGAAAGCGGTGGCCACCGACTGGGGG 1140
QY 1141 TCCGCTTACGACGCGCCGACGCTGTGCGCTGACTCGGCGCGCTGCTTACGCGGAC 1200
Db 1141 TCCCGTTCACGACGCGCCGACGCTGTGCGCTGACTCGGCGCGCTGCTTACGCGGAC 1200
QY 1201 ACACGTGNGAGTAAAGGAAGCATGATG 1229
Db 1201 ACACGTGNGAGTAAAGGAAGCATGATG 1229
RESULT 5
A87609 1235 bp DNA PAT 22-JAN-2000
LOCUS Sequence 6 from Patent WO9836089.
ACCESSION A87609
VERSION A87609.1 GI:6736249
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1235)
AUTHORS Flohe, L. and Singh, M.
TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL Patent: WO 9836089-A 20-AUG-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
LOCATION/Qualifiers
1..1235
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 235 a 395 c 384 g 220 t 1 others
ORIGIN

Query Match 99.7% Score 1225.4; DB 5; Length 1235;
Best Local Similarity 99.8%; Pred. No. 6.4e-167;
Matches 1226; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACTTGCAGATTAAATGAACTTCTTACACGTAAGCTGTAAGTATCGAAGGGGTATC 60
Db 1 ACTTGCAGATTAAATGAACTTCTTACACGTAAGCTGTAAGTATCGAAGGGGTATC 60
QY 61 ATGCGGCTGCTATTCGACGACGACAAACAAACAAAGTTCGGGTGGCCATCAACCCG 120
Db 61 ATGCGGCTGCTATTCGACGACGACAAACAAACAAAGTTCGGGTGGCCATCAACCCG 120
QY 121 GCGGCGTCCGGAATTAACCCGCTGCTGCGCATGAGTGTCTATCCAGCAGGTGCCGA 180
Db 121 GCGGCGTCCGGAATTAACCCGCTGCTGCGCATGAGTGTCTATCCAGCAGGTGCCGA 180
QY 181 GAGGGCTGGCTATCAGCAGCGGATTTCAAGGGGCGGCGGCACTGGTGGGAC 240
Db 181 GAGGGCTGGCTATCAGCAGCGGATTTCAAGGGGCGGCGGCACTGGTGGGAC 240
QY 241 GCGGACCAAGTGTGGGCGACGCTGATTTATGCTCAAGTCAAAAGACGATAGCGGG 300
Db 241 GCGGACCAAGTGTGGGCGACGCTGATTTATGCTCAAGTCAAAAGACGATAGCGGG 300
QY 301 GAATAGGCGCGCTGGGACACGCGGATCTGTGTCAGCTTGTGCAATTTGGCGGCTAC 360
Db 301 GAATAGGCGCGCTGGGACACGCGGATCTGTGTCAGCTTGTGCAATTTGGCGGCTAC 360
QY 361 GAGCTTGCACCGATGCTTGTGATTCGGGACGACGTAATTTGCTTGCAGACGCTGC 420
Db 361 GAGCTTGCACCGATGCTTGTGATTCGGGACGACGTAATTTGCTTGCAGACGCTGC 420
QY 421 AGACGCGCAAGCGCACTACCCCTCTTCCGCGATGAGCGAAGTGGCGGCTGACTCG 480
Db 421 AGACGCGCAAGCGCACTACCCCTCTTCCGCGATGAGCGAAGTGGCGGCTGACTCG 480
QY 481 CCGCCCAAGTGTGGCTTACACATGATGGAACCCAAAGGGGGCGCGGTGTCTATGG 540
Db 481 CCGCCCAAGTGTGGCTTACACATGATGGAACCCAAAGGGGGCGCGGTGTCTATGG 540

QY 541 GCGGGGTGCGCGGCTGCAACCGCGGACGTCGTGTGATTCGGCGCGGACCGCGGCT 600
Db 541 GCGGGGTGCGCGGCTGCAACCGCGGACGTCGTGTGATTCGGCGCGGACCGCGGCT 600
QY 601 ACACGACGCGCGCATTCGCAACGACGATGCGGCGGACGCTTACGATCATCA 660
Db 601 ACACGACGCGCGCATTCGCAACGACGATGCGGCGGACGCTTACGATCATCA 660
QY 661 TCGACAACTTGGCACTCGACGCGGAGTTCTGCGCGGATCCACACGCTACTCAT 720
Db 661 TCGACAACTTGGCACTCGACGCGGAGTTCTGCGCGGATCCACACGCTACTCAT 720
QY 721 CGGCTTACGAGCTCGAGGCTGCGGTCGAACGTCGCGGACCTGCTGATTTGGGCGCTCGG 780
Db 721 CGGCTTACGAGCTCGAGGCTGCGGTCGAACGTCGCGGACCTGCTGATTTGGGCGCTCGG 780
QY 781 TCGCAGGCGCCCAAGGACACCAATTAATGCTGCAATTCATTTGCGGATATGAACAG 840
Db 781 TCGCAGGCGCCCAAGGACACCAATTAATGCTGCAATTCATTTGCGGATATGAACAG 840
QY 841 GTGCGCTACTGTGATATATGATCATGACACGCGGCTGTTTGAAGGCTCAGACCGA 900
Db 841 GTGCGCTACTGTGATATATGATCATGACACGCGGCTGTTTGAAGGCTCAGACCGA 900
QY 901 CCACCTAGACACCGCGGACGTCGCGGTCGACAGACGCTGTTTACTGCGGCGGACGA 960
Db 901 CCACCTAGACACCGCGGACGTCGCGGTCGACAGACGCTGTTTACTGCGGCGGACGA 960
QY 961 TCCCGCGCTGCTGCGGACGAGAGTGCACCTACGCGCTGACCAACGACGATGCCGTATG 1020
Db 961 TCCCGCGCTGCTGCGGACGAGAGTGCACCTACGCGCTGACCAACGACGATGCCGTATG 1020
QY 1021 TCTGAGAGTTTCCGACCATGCTGTGCGGCGGCGCTGCGGCTGCAATCGGCACTAGCCA 1080
Db 1021 TCTGAGAGTTTCCGACCATGCTGTGCGGCGGCGCTGCGGCTGCAATCGGCACTAGCCA 1080
QY 1081 AAGGTCTTTCGACGACGAGTGGGCTTACTGTCCGAACGGGTG :CAACGACCTGGGG 1140
Db 1081 AAGGTCTTTCGACGACGAGTGGGCTTACTGTCCGAACGGGTG :CAACGACCTGGGG 1140
QY 1141 TCCGCTTACCGAGCGCGCAACGCTGCTGCGCTGACTGCGGCGGCTGCTTACGCGGACG 1200
Db 1141 TCCGCTTACCGAGCGCGCAACGCTGCTGCGCTGACTGCGGCGGCTGCTTACGCGGACG 1200
QY 1201 ACACGTGNGAGTAAAGGAAGCATGATG 1229
Db 1201 ACACGTGNGAGTAAAGGAAGCATGATG 1229
RESULT 6
A89749 1235 bp DNA PAT 22-JAN-2000
LOCUS Sequence 6 from Patent WO9832862.
ACCESSION A89749
VERSION A89749.1 GI:6738283
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1235)
AUTHORS Flohe, L. and Singh, M.
TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
JOURNAL Patent: WO 9832862-A 30-JUL-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
LOCATION/Qualifiers
1..1235
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 235 a 395 c 384 g 220 t 1 others
ORIGIN

Query Match 99.7%; Score 1225.4; DB 5; Length 1235;
 Best Local Similarity 99.8%; Pred. No. 6.4e-167;
 Matches 1226; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 ATCTTGACAGTTATGAACTTCTTACACGTAAGCGTATGACGAGAGGGGTATC 60
   |||||||
Db 1 ATCTTGACAGTTATGAACTTCTTACACGTAAGCGTATGACGAGAGGGGTATC 60
QY 61 ATGCGGCTGCGGTATTCGACGACGACCAAAACAAAGAAATTCGGGGTGGCATCACC 120
   |||||||
Db 61 ATGCGGCTGCGGTATTCGACGACGACCAAAACAAAGAAATTCGGGGTGGCATCACC 120
QY 121 GCGGCGGTGGGGAATAACCCGTCCTGGCCATGAGTGTCTATCCAGCAGGTGCGGA 180
   |||||||
Db 121 GCGGCGGTGGGGAATAACCCGTCCTGGCCATGAGTGTCTATCCAGCAGGTGCGGA 180
QY 181 GAGGCGTCCGCTATCACGACGACGCGATTTCAAGGCGGACGCGCAATGGTCGAC 240
   |||||||
Db 181 GAGGCGTCCGCTATCACGACGACGCGATTTCAAGGCGGACGCGCAATGGTCGAC 240
QY 241 GCGGACAGGTGCGGCGGACGCTGATTTATGCTCAAGAGTCAAGAACCGATAGCG 300
   |||||||
Db 241 GCGGACAGGTGCGGCGGACGCTGATTTATGCTCAAGAGTCAAGAACCGATAGCG 300
QY 301 GAATACGCGCGCTCGGACGACGCGGATCTGTTACGTTCTTGGCCGCGCTAC 360
   |||||||
Db 301 GAATACGCGCGCTCGGACGACGCGGATCTGTTACGTTCTTGGCCGCGCTAC 360
QY 361 GTGCTTGACACCGATGCGTGTGATTCGCGACACGTCGTAATGGCTAGAGACCG 420
   |||||||
Db 361 GTGCTTGACACCGATGCGTGTGATTCGCGACACGTCGTAATGGCTAGAGACCG 420
QY 421 AGACGCGGAGGAGGACGACTACCCCTGCTGGCCGATGAGCGAAGTGGCGGTCGAC 480
   |||||||
Db 421 AGACGCGGAGGAGGACGACTACCCCTGCTGGCCGATGAGCGAAGTGGCGGTCGAC 480
QY 481 CGGCGGAGGTGGCGCTTACACACCTGATGCAAGCCAGGCGCGGTGTCTGATG 540
   |||||||
Db 481 CGGCGGAGGTGGCGCTTACACACCTGATGCAAGCCAGGCGCGGTGTCTGATG 540
QY 541 GGGGGGTGGCGGCTGCAACCGCGGAGTGTGATTCGGCGCGGCGCGCGCGGT 600
   |||||||
Db 541 GGGGGGTGGCGGCTGCAACCGCGGAGTGTGATTCGGCGCGGCGCGCGCGGT 600
QY 601 ACAACGACCGCGCATCGCCACAGCGGATGGGCGGACGCTTACGTTAGACATCA 660
   |||||||
Db 601 ACAACGACCGCGCATCGCCACAGCGGATGGGCGGACGCTTACGTTAGACATCA 660
QY 661 TCGACAAATTTGGGCACTCGACGCGGATTTGGGGCGGATCCACTCGCTACT 720
   |||||||
Db 661 TCGACAAATTTGGGCACTCGACGCGGATTTGGGGCGGATCCACTCGCTACT 720
QY 721 CGGCGTACGAGCTCGAGGCTGCGTCAACGTCGCGACCTGATGATGGGGCGG 780
   |||||||
Db 721 CGGCGTACGAGCTCGAGGCTGCGTCAACGTCGCGACCTGATGATGGGGCGG 780
QY 781 TCGCAGGCGCCAAAGCACCCTAATTAATGTCGATTCGATTCGCGCATATGA 840
   |||||||
Db 781 TCGCAGGCGCCAAAGCACCCTAATTAATGTCGATTCGATTCGCGCATATGA 840
QY 841 GTCGCGTACGAGTATGATAGCATGACGACGAGGGGCGGCTTCAAGGCTACAC 900
   |||||||
Db 841 GTCGCGTACGAGTATGATAGCATGACGACGAGGGGCGGCTTCAAGGCTACAC 900
QY 901 CCACCTACGACACCGGAGCTGCGCGTGCACGACGCTGTTTACGCGTGGGGA 960
   |||||||
Db 901 CCACCTACGACACCGGAGCTGCGCGTGCACGACGCTGTTTACGCGTGGGGA 960
QY 961 TCGCGGCGCTGCGTGGGGAAGAGTGCATAGCGGCTGACCAACGCGAGATG 1020
   |||||||
Db 961 TCGCGGCGCTGCGTGGGGAAGAGTGCATAGCGGCTGACCAACGCGAGATG 1020
QY 1021 TCGTGTAGCTTCCGACGATGCTGCGGCGGCGGTGCGGTCGAATCCGGCAT 1080

```

```

Db 1021 TCGTGTAGCTTCCGACGATGCTGCGGCGGCGGTGCGGTCGAATCCGGCAT 1080
QY 1081 AAGGCTTTTCGACGACGACGAGGGGCTTACTGTCCGAACGGGTGGCCACGACT 1140
   |||||||
Db 1081 AAGGCTTTTCGACGACGACGAGGGGCTTACTGTCCGAACGGGTGGCCACGACT 1140
QY 1141 TCGCGTTCGAGGAGGCGGCGGCGTGGCTGACCTCGGCGGCTTACGCGAGC 1200
   |||||||
Db 1141 TCGCGTTCGAGGAGGCGGCGGCGTGGCTGACCTCGGCGGCTTACGCGAGC 1200
QY 1201 ACAGCTCGGAGTAAAGGACGATGATG 1229
   |||||||
Db 1201 ACAGCTCGGAGTAAAGGACGATGATG 1229

```

```

RESULT 7
A87613
LOCUS A87613 1236 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 10 from Patent WO9836089.
ACCESSION A87613
VERSION A87613.1 GI:6736253
KEYWORDS
SOURCE
ORANISM unidentified.
REFERENCE 1 (bases 1 to 1236)
AUTHORS Flohe, J., and Singh, M.
TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL Patent: WO 9836089-A-20-AUG-1998;
FLOHE LEOPOLD (DE); SINGH MANAVIR (DE)
FEATURES
Source location/Qualifiers
1..1236
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 236 a 395 c 385 g 220 t
ORIGIN

```

Query Match 98.9%; Score 1215.4; DB 5; Length 1236;
 Best Local Similarity 99.8%; Pred. No. 1.7e-165;
 Matches 1227; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```

QY 1 ATCTTGACAGTTATGAACTTCTTACACGTAAGCGTATGACGAGAGGGGTATC 60
   |||||||
Db 1 ATCTTGACAGTTATGAACTTCTTACACGTAAGCGTATGACGAGAGGGGTATC 60
QY 61 ATGCGGCTGCGGTATTCGACGACGACCAAAACAAAGAAATTCGGGGTGGCATCACC 120
   |||||||
Db 61 ATGCGGCTGCGGTATTCGACGACGACCAAAACAAAGAAATTCGGGGTGGCATCACC 120
QY 121 GCGGCGGTGGGGAATAACCCGTCCTGGCCATGAGTGTCTATCCAGCAGGTGCGGA 180
   |||||||
Db 121 GCGGCGGTGGGGAATAACCCGTCCTGGCCATGAGTGTCTATCCAGCAGGTGCGGA 180
QY 181 GAGGCGTCCGCTATCACGACGACGCGATTTCAAGGCGGACGCGCAATGGTGGG 240
   |||||||
Db 181 GAGGCGTCCGCTATCACGACGACGCGATTTCAAGGCGGACGCGCAATGGTGGG 240
QY 241 GCGGACAGGTGCGGCGGACGCTGATTTATGCTCAAGGTCGCAAGAACCGATAGCG 300
   |||||||
Db 241 GCGGACAGGTGCGGCGGACGCTGATTTATGCTCAAGGTCGCAAGAACCGATAGCG 300
QY 301 GAATACGCGCGCTCGGACGACGCGGCTGTTTACGTTCTTGGCCGCGCTAC 359
   |||||||
Db 301 GAATACGCGCGCTCGGACGACGCGGCTGTTTACGTTCTTGGCCGCGCTAC 359
QY 360 GTCGCTTGACACCGATGCGTGTGATTCGCGGACGCTGATTAATGGCTAGAGACCG 419
   |||||||
Db 360 GTCGCTTGACACCGATGCGTGTGATTCGCGGACGCTGATTAATGGCTAGAGACCG 419
QY 420 CAGACCGCGGAGGCGGACGCTACCCCTGCTGGCGATGAGCGAAGTGGCGGTCGACT 479
   |||||||

```



```

Db 421 CAGACCCGCGACGCGCTACCCCTGCTTGCCCGCATGAGCGAACTCCGCGCTGCACTC 480
Oy 480 GCGCGCCAGGTTGGGCTTACACGATGCGAACCCAGGCGCGCGCTGTCGATG 539
Db 481 GCGCGCCAGGTTGGGCTTACACGATGCGAACCCAGGCGCGCGCTGTCGATG 540
Oy 540 GCGCGCGCTGCGCGCTGCGAACCGCGCGCGCTGTCGATGCGCGCGCGCGCGCG 599
Db 541 GCGCGCGCTGCGCGCGCTGCGAACCGCGCGCGCTGTCGATGCGCGCGCGCGCG 600
Oy 600 TACACGCGACCGCGCTGCGAACCGCGCGCGCTGTCGATGCGCGCGCGCGCG 659
Db 601 TACACGCGACCGCGCTGCGAACCGCGCGCGCTGTCGATGCGCGCGCGCGCG 660
Oy 660 ATGACAACTTGGGCACTGCGAACCGCGCGCGCTGTCGATGCGCGCGCGCGCG 719
Db 661 ATGACAACTTGGGCACTGCGAACCGCGCGCGCTGTCGATGCGCGCGCGCGCG 720
Oy 720 TCGGCTTACGAGCTGCGAGGCTGCGTCAACGCTGCGCGCGCTGTCGATGCGCG 779
Db 721 TCGGCTTACGAGCTGCGAGGCTGCGTCAACGCTGCGCGCGCTGTCGATGCGCG 780
Oy 780 GTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 839
Db 781 GTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
Oy 840 GTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 899
Db 841 GTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Oy 900 ACCACCTAGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 959
Db 901 ACCACCTAGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
Oy 960 ATGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1019
Db 961 ATGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Oy 1020 GTGCTGAGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1079
Db 1021 GTGCTGAGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
Oy 1080 AAAGGCTTTCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1139
Db 1081 AAAGGCTTTCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
Oy 1140 GTGCGCTTTCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1199
Db 1141 GTGCGCTTTCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1200
Oy 1200 CACACGCTGCGGAGTAAAGGAGGATGATG 1229
Db 1201 CACACGCTGCGGAGTAAAGGAGGATGATG 1230

```

RESULT 8

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1236 bp DNA
Sequence 10 from Patent WO9832862.
A89753
A89753.1 GI:6738287

PAT 22-JAN-2000

unclassified.
unclassified.

1 (bases 1 to 1236)
F.lobae, L. and Singh, M.

L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
Patent: WO 9832862-A 30-JUL-1998;

FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
Location/Qualifiers

1..1236
/organism="unclassified"

BASE COUNT 236 a 395 c 385 g 220 t
ORIGIN

Query Match 98.9%; Score 1215.4; DB 5; Length 1236;
Best Local Similarity 99.8%; Pred. No. 1.7e-165;

Matches 1227; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```

Oy 1 ATCTTGCAAGTTAATCGAATCTTCTTACACTGAAAGCGTACAGTATCGAAGGGTAAATC 60
Db 1 ATCTTGCAAGTTAATCGAATCTTCTTACACTGAAAGCGTACAGTATCGAAGGGTAAATC 60
Oy 61 ATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db 61 ATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Oy 121 GCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Db 121 GCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Oy 181 GAGGCTGCGCTATACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Db 181 GAGGCTGCGCTATACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Oy 241 GCGGACGAGTGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Db 241 GCGGACGAGTGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Oy 301 GAATACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 359
Db 301 GAATACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Oy 360 CGTGTGACCGATGCGTGTGATTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 419
Db 361 CGTGTGACCGATGCGTGTGATTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
Oy 420 CAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 479
Db 421 CAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
Oy 480 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 539
Db 481 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Oy 540 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 599
Db 541 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Oy 600 TACAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 659
Db 601 TACAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
Oy 660 ATGACAACTTGGGCACTGCGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 719
Db 661 ATGACAACTTGGGCACTGCGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
Oy 720 TCGGCTTACGAGCTGCGAGGCTGCGTCAACGCTGCGCGCGCGCGCGCGCG 779
Db 721 TCGGCTTACGAGCTGCGAGGCTGCGTCAACGCTGCGCGCGCGCGCGCGCG 780
Oy 780 GTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 839
Db 781 GTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
Oy 840 GTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 899
Db 841 GTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Oy 900 ACCACCTAGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 959
Db 901 ACCACCTAGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960

```

```

QY 960 ATGCGCGCTGGTGGCCGAGAGAGCTGACCTACGCGCTGACCAAGCGAGCATGCGTAT 1019
DB 961 ATGCGCGCTGGTGGCCGAGAGAGCTGACCTACGCGCTGACCAAGCGAGCATGCGTAT 1020
QY 1020 GTGCTGACGCTTGGCCGAGAGAGCTGACCTACGCGCTGACCAAGCGAGCATGCGTAT 1079
DB 1021 GTGCTGACGCTTGGCCGAGAGAGCTGACCTACGCGCTGACCAAGCGAGCATGCGTAT 1080
QY 1080 AAGGCTCTTCCAGCAGCAGAGAGAGCTGACCTACGCGCTGACCAAGCGAGCATGCGTAT 1139
DB 1081 AAGGCTCTTCCAGCAGCAGAGAGAGCTGACCTACGCGCTGACCAAGCGAGCATGCGTAT 1140
QY 1140 GTGCGCTTACGAGAGAGAGAGCTGACCTACGCGCTGACCAAGCGAGCATGCGTAT 1199
DB 1141 GTGCGCTTACGAGAGAGAGAGCTGACCTACGCGCTGACCAAGCGAGCATGCGTAT 1200
QY 1200 CACAGCTGAGAGAGAGAGAGCTGACCTACGCGCTGACCAAGCGAGCATGATG 1229
DB 1201 CACAGCTGAGAGAGAGAGAGCTGACCTACGCGCTGACCAAGCGAGCATGATG 1230

RESULT 9
LOCUS A87607 1237 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 4 from Patent WO9836089.
ACCESSION A87607
VERSION A87607.1 GI:6736247
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1237)
AUTHORS Flohe, L. and Singh, M.
TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL Patent: WO 9836089-A 20-AUG-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
source 1..1237
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 236 a 394 c 386 g 221 t
ORIGIN
Query Match 98.8%; Score 1213.8; DB 5; Length 1237;
Best Local Similarity 99.7%; Pred. No. 2.9e-165;
Matches 1226; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 1 ATCTGAGATTATGAGACTTCTCACACTGAGAGCTGACGATGATGAGAGGGTATC 60
DB 1 ATCTGAGATTATGAGACTTCTCACACTGAGAGCTGACGATGATGAGAGGGTATC 60
QY 61 ATGCGGCTGGATATTCGACGAGAGCAAAACAGAGATTCGGGTGCGCATACCCCG 120
DB 61 ATGCGGCTGGATATTCGACGAGAGCAAAACAGAGATTCGGGTGCGCATACCCCG 120
QY 121 GCGCGGCTGGAGAGCAAAACAGAGATTCGGGTGCGCATACCCCG 180
DB 121 GCGCGGCTGGAGAGCAAAACAGAGATTCGGGTGCGCATACCCCG 180
QY 181 GAGGCTGGCTATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
DB 181 GAGGCTGGCTATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 241 GCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 241 GCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 301 GAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 359
DB 301 GAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 360 CGTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 419

```

```

DB 361 CGTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 420 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 479
DB 421 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 480 GCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 539
DB 481 GCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 540 GCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 599
DB 541 GCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 600 TACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 659
DB 601 TACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 660 ATGCAAAACTTGGCAACTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 719
DB 661 ATGCAAAACTTGGCAACTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 720 TCGGCTTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 779
DB 721 TCGGCTTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 780 GTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 839
DB 781 GTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 840 GTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 899
DB 841 GTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 900 AACAGCTTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 959
DB 901 AACAGCTTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 960 ATGCGGCTGGTGGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1019
DB 961 ATGCGGCTGGTGGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1020 GTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1079
DB 1021 GTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1080 AAGGCTCTTCCAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1139
DB 1081 AAGGCTCTTCCAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 1140 GTGCGCTTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1199
DB 1141 GTGCGCTTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 1200 CACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1229
DB 1201 CACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1230

RESULT 10
LOCUS A89747 1237 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 4 from Patent WO9832862.
ACCESSION A89747
VERSION A89747.1 GI:6738281
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1237)
AUTHORS Flohe, L. and Singh, M.
TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM

```

JOURNAL Patent: WO 9832862-A 30-JUL-1998;

FLOHE LEBOLD (DE); SINGH MAHAVIR (DE)

FEATURES

Location/Qualifiers

source

1. 1237

/organism="unidentified"

/db_xref="taxon:32644"

BASE COUNT

236 a 394 c 386 g 221 t

ORIGIN

Query Match 98.8%; Score 1213.8; DB 5; Length 1237;

Best Local Similarity 99.7%; Pred. No. 2.9e-165;

Matches 1226; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 ATCTTCAGATTAAATGAACTTTCTTACACGTAAGCTGATACAGTATCGAGAGGGTAAATC 60
 Db 1 ATCTTCAGATTAAATGAACTTTCTTACACGTAAGCTGATACAGTATCGAGAGGGTAAATC 60
 QY 61 ATGGCGGTGGTATTTCGACCGGACCAAAACAGCAATTCGGGTGGCCATACCCCG 120
 Db 61 ATGGCGGTGGTATTTCGACCGGACCAAAACAGCAATTCGGGTGGCCATACCCCG 120
 QY 121 GCGGCGGTGCGGAACTAACCCGTCGTGGCCATAGAGTCTCATCCAGCGAGTGGCGGA 180
 Db 121 GCGGCGGTGCGGAACTAACCCGTCGTGGCCATAGAGTCTCATCCAGCGAGTGGCGGA 180
 QY 181 GAGGCGTGGCTATATCACGACCGGATTTCAAGGCGGACGCGGCACTGTCGCGACC 240
 Db 181 GAGGCGTGGCTATATCACGACCGGATTTCAAGGCGGACGCGGCACTGTCGCGACC 240
 QY 241 GCGGACGAGTGGGCGGACGCTGATTTATTCGTAAGGTCAAAAGACGATGCGGGG 300
 Db 241 GCGGACGAGTGGGCGGACGCTGATTTATTCGTAAGGTCAAAAGACGATGCGGGG 300
 QY 301 GAATACGCGCGCTGCGACACGCGGCGATCTTGTTCATTCGATTTGGCGCGCTCA 359
 Db 301 GAATACGCGCGCTGCGACACGCGGCGATCTTGTTCATTCGATTTGGCGCGCTCA 359
 QY 360 CGTCTTGCACGAGTGGCTTTGGATTCCGCGACACGTCATTTGCCATTCGACAGCCGTC 419
 Db 360 CGTCTTGCACGAGTGGCTTTGGATTCCGCGACACGTCATTTGCCATTCGACAGCCGTC 419
 QY 420 CAGACCGCGGAGGCGGACGCTGATTTATTCGTAAGGTCAAAAGACGATGCGGGG 479
 Db 420 CAGACCGCGGAGGCGGACGCTGATTTATTCGTAAGGTCAAAAGACGATGCGGGG 479
 QY 480 GCGGCGGAGTGGGCGGCTTACCACTGATGCGAACCAGGCGGCGGCTGTGTGATG 539
 Db 480 GCGGCGGAGTGGGCGGCTTACCACTGATGCGAACCAGGCGGCGGCTGTGTGATG 539
 QY 540 GCGGCGGAGTGGGCGGCTTACCACTGATGCGAACCAGGCGGCGGCTGTGTGATG 599
 Db 540 GCGGCGGAGTGGGCGGCTTACCACTGATGCGAACCAGGCGGCGGCTGTGTGATG 599
 QY 600 TACAACGCGCGGCTGCGACGCGGATGGGCGGACGCTTACGATGATGATGATGATG 659
 Db 600 TACAACGCGCGGCTGCGACGCGGATGGGCGGACGCTTACGATGATGATGATGATG 659
 QY 660 ATGACAAATCTTGGCACTGCAAGCGGAGTGTGGCGGAGTCCACACTGCTACTCA 719
 Db 660 ATGACAAATCTTGGCACTGCAAGCGGAGTGTGGCGGAGTCCACACTGCTACTCA 719
 QY 720 TGGGCGTACGAGTGGGCGGCTTACCACTGATGCGAACCAGGCGGCGGCTGTGTG 779
 Db 720 TGGGCGTACGAGTGGGCGGCTTACCACTGATGCGAACCAGGCGGCGGCTGTGTG 779
 QY 780 GTGCGAGGCGGCGGACGCGGATGATGATGATGATGATGATGATGATGATGATG 839
 Db 780 GTGCGAGGCGGCGGACGCGGATGATGATGATGATGATGATGATGATGATGATG 839
 QY 840 GGTGGCGTACGAGTGGGCGGCTTACCACTGATGCGAACCAGGCGGCGGCTGTGTG 899
 Db 840 GGTGGCGTACGAGTGGGCGGCTTACCACTGATGCGAACCAGGCGGCGGCTGTGTG 899

QY 900 ACCACTTACGACCCGACGCTTCCGTCGACGACGACGCTTTTACTGCTGGCGAAC 959
 Db 901 ACCACTTACGACCCGACGCTTCCGTCGACGACGACGCTTTTACTGCTGGCGAAC 960
 QY 960 ATGCCCGGCTGGTGGCGGAGGCTGACCTGCGGCGGAGTACCAACGACGATGGCGAT 1019
 Db 961 ATGCCCGGCTGGTGGCGGAGGCTGACCTGCGGCGGAGTACCAACGACGATGGCGAT 1020
 QY 1020 GTGCTGAGCTTCCGACCATGGCTGGGCGGCGGCGGCTGCAATCCGCACTACCC 1079
 Db 1021 GTGCTGAGCTTCCGACCATGGCTGGGCGGCGGCGGCTGCAATCCGCACTACCC 1080
 QY 1080 AAGGCTTTTGCACGACGAGGCGGCTTACTGTCCGAAGGGGTGCCACGACCTGGGG 1139
 Db 1081 AAGGCTTTTGCACGACGAGGCGGCTTACTGTCCGAAGGGGTGCCACGACCTGGGG 1140
 QY 1140 GTGCGTTCACGACCGGCGGCGGCTGCTGGCTGACCTGCGGCGGCTGATGCGCGAG 1199
 Db 1141 GTGCGTTCACGACCGGCGGCGGCTGCTGGCTGACCTGCGGCGGCTGATGCGCGAG 1200
 QY 1200 CACAGCTCNGAGTAAGGGAAGCGATGATG 1229
 Db 1201 CACAGCTCNGAGTAAGGGAAGCGATGATG 1230

RESULT 11

MTV002

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REMARK

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REMARK

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REMARK

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REMARK

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REMARK

Cole, S.T., Broesch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eigmeier, K., Gas, S., Barry III, C.E., Tekle, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltham, J., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, R., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, S., Osborne, J., Quail, M.A., Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skelton, S., Squares, S., Squires, R., Sultson, J.E., Taylor, K., Whitehead, S., and Barrall, B.G.
 Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence
 Nature 393 (6685), 537-544 (1998)
 Erratum: [[published erratum appears in Nature 1998 Nov 12:396(6707):1901]]
 2 (bases 1 to 56414)
 Parkhill, J.
 Direct Submission
 Submitted (11-JUN-1998) submitted on behalf of the *Mycobacterium tuberculosis* sequencing and mapping teams, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk
 On Jun 27, 1998 this sequence version replaced gi:2624256.
 Notes:
 Details of *M. tuberculosis* sequencing at the Sanger Centre are available on the World Wide Web.
 (URL, http://www.sanger.ac.uk/Projects/M_tuberculosis/) CDS have been renumbered from the original cosmid submissions to the old gene designations are in brackets after the new gene numbers.
 Gene prediction was based on a Hidden Markov Model of 18 genes implemented in TParse (Krogh) supplemented with visual inspection of positional base preference in codons, especially where there is an increase in the observed/expected third position G + C.

CAUTION: In some cases we may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, or ttg) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

FEATURES

```

source
    1..56414
    /organism="Mycobacterium tuberculosis"
    /strain="H37Rv"
    /db_xref="taxon:1773"
    <1..>56414
    /note="fragment designated V002. Does not represent a physical clone"
    <1..233
    /organism="Mycobacterium tuberculosis"
    /strain="H37Rv"
    /db_xref="taxon:1773"
    /clone="Y154"
    /complement(3..527)
    /gene="recX"
    /complement(3..527)
    /gene="recX"
    /note="Rv2736c, (MTV002.01c), len: 174 aa. recX, similar to eg. RECX_PSEAE_P37860 regulatory protein recX from Pseudomonas aeruginosa (153 aa), fasta scores; opt: 161 z-score: 257.2 E(): 3.6e-07, 30.7% identity in 137 aa overlap. Overlaps and extends CDS from overlapping cosmid MTCY154.16c"
    /codon_start=1
    /transl_table=11
    /product="recX"
    /protein_id="CA15533.1"
    /db_xref="GI:2624257"
    /transl_table="MTVSCPSPSTSEPEQARACIRLTARSRRAELAQAKRCY PEDIGNRYLDLAAGLYDTDFEOMVQSRANAASKRALAEHAKGVDDVITTVLGGIDACGERAEKTLRARLRREVLDIDGTDEARVSRRLVAMLRGGYQTLACEY VIAELAEERERRRV"
    /complement(493..2865)
    /gene="recA"
    /complement(493..2865)
    /gene="recA"
    /complement(493..2865)
    /note="Rv2737c, (MTV002.02c), len: 790 aa. recA, identical to REC_A_MYCYU_P26345 recA protein (790 aa). Contains self-splicing protein element (intein) from 2294 to 969 (C). similar to intein II from TR:E332317 (EMBL:Y13030) DNA-directed DNA polymerase (EC 2.7.7.7) from Thermococcus sp. (1829 aa), fasta scores; opt: 81 z-score: 235.2 E(): 6e-06, 24.6% identity in 183 aa overlap. Contains PS00017 ATP/GTP-binding site motif A (P-loop), PS00321 recA signature, and PS00881 protein splicing signature. See Davis et al, (1992) Cell 71(2):201-210"
    /codon_start=1
    /transl_table=11
    /product="recA"
    /protein_id="CA15533.1"
    /db_xref="GI:2624258"
    /db_xref="SWISS-PROT:P26345"
    /translation="MTOPPRKALELAVAQIEKSYGKSVMLRGDEARODISYIPG SIADVAIGIGLPGRIYETGPESSSKTTVAHANNANNAAGVAAFIDAEHALDP DYAKLGYDTSLVSOPDTGEQALEIDMLIRSGALDIVDSVAALVPAELEGEG GDHVGQLARLMSQALRKMTGALNNGTALFINGLRDRIQVFGSPETTGKALF YASVMDVAREVETLKDGINAVGNRTRVVKNKCLAEGRIFDVTGTHRIEYVVG RKPPIHVAAADGTLHAPVYVSMFDGRTDVIYIGIAGATVMAIPDHKVIIEYVGMRA AGELRKGRVAAQPRPFGDSAPIPADHARLIGLIGDGDGNGVGGKTPINFINVOR ALIDYPRATATTCGAAPGCRSLATIHAPGNGVADICCOAGIYCKLAKWETIYN WPFEDTADIVGNLIRGLERSDQWVSBEQTAGLRVGTITTSQDLAHIDMLIRFVY GSTVADYPTQKRPSIVNGRIQSKRQVEFVRISGMONTVAAGASVPMGRGALKY AIPETQGRRGSOATYLAEMTDAYLVNIDERGVTQEEAAMTGAASGDRGCKQV LGASVLRDRVOALADLDKFLDMLAEELRYSVIREVLTARRAFDDEVEELHTL VAEGVYVNGSPFPOAEFDILYGGISFEELSLIDMGDOGLIRKSGMWFYEEQDQ GQKEARAFLEVNADVDEIKIKIKIKICIGAVVTTDDPSNGVLPAPVDF"

```

```

misc_feature
    complement(790..2112)
    /gene="recA"
    /note="recA intein"
    /complement(790..807)
    /gene="recA"
    /note="PS00881 Protein splicing signature"
    /complement(2197..2223)
    /gene="recA"
    /note="PS00321 recA signature"
    /complement(2644..2667)
    /gene="recA"
    /note="PS00017 ATP/GTP-binding site motif A (P-loop)"
    /complement(2873..2877)
    /note="possible RBS upstream of Rv2737c"
    /complement(3247..3453)
    /gene="Rv2738c"
    /complement(3247..3453)
    /note="Rv2738c, (MTV002.03c), len: 68 aa. Unknown, N-terminus is highly similar to the N-terminus of the upstream ORF MTV002.07c (78.4% identity in 37 aa overlap); also similar to AL020958|SC4H8.5 Streptomyces coelicolor cosmid 4H8 (64aa) opt: 185 z-score: 283.5 E(): 2.9e-08; 39.7% identity in 63 aa overlap"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein Rv2738c"
    /protein_id="CA15534.1"
    /db_xref="GI:2624260"
    /db_xref="SPTREMBL:033281"
    /translation="MLAGVRLTFEHERVALHFGAAYGSVLDHYLTGFDGRSAQAI EDGVEPRDVRAICADFDVPHDRW"
    /complement(3464..4630)
    /gene="Rv2739c"
    /complement(3464..4630)
    /note="Rv2739c, (MTV002.04c), len: 388 aa. Probable transferase, similar to eg. TR:Q51560 (EMBL:L28170) Pseudomonas aeruginosa ribonuclease (426 aa), fasta scores; opt: 178 z-score: 226.3 E(): 1.9e-05, 25.9% identity in 425 aa overlap. Equivalent to Mycobacterium leprae protein MLCB33.02c (392 aa); fasta scores gpl294723|MLCB33.2 Mycobacterium leprae cosmid B33 opt: 2112 z-score: 2364.5 E(): 0; 80.9% identity in 388 aa overlap"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein Rv2739c"
    /protein_id="CA15535.1"
    /db_xref="GI:2624261"
    /db_xref="SPTREMBL:033282"
    /translation="MRAVAVAGPDPGSEFAIALCQFRRAADPTLTGVEETEAR AACIDVNEIDGLATDRDLDAGARIRRAQNAVLVPRRLDAETLVSDVITRAGG MAEELGIPWELNPHPLVPSKGLPPIGSLAAGVIRKRLDAPRALVTSWRAG LRQRAARVVEIIGLPARDPGLRLIATLPALVPRDMPAPAAVVGPLEPDRVLA IPAGTGVVVAAPSTALGTAGLEVALOSTVGETVPSGSLVYSRLSGADLTVPV AVAGLSQAEILTRADIVICGGHGWAKLLGVGMVYVPGGDQMEIANTRVRGGS AVLRPLTADALVAANEVLSPRFRRAARAASVGAADPVRVCHDALALAG"
    /complement(4674..5123)
    /gene="Rv2740"
    /complement(4674..5123)
    /note="Rv2740, (MTV002.05), len: 149 aa. unknown, similar to M. leprae hypothetical protein TR:Q49850 (EMBL:294723) MLCB33.03 (B2235.F3.140) (178 aa), fasta scores; opt: 498 z-score: 538.9 E(): 7.3e-178, 51.6% identity in 161 aa overlap"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein Rv2740"
    /protein_id="CA15536.1"
    /db_xref="GI:2624262"
    /db_xref="SPTREMBL:033283"

```

gene
CDS

/translation="MAETLSPENPETTALRAVEALNALONEDPTVVALGDL
VYENVGRIRGRRATILRMGRVFEVIRHIGDGAVALTEPDAIILGPIRV
QFWVGVEVDGRTILMRDIFDYDMFKGLRLGLVALVPSLATL-
5355..6932
/gene="PE_PGRS"
5355..6932
/gene="PE_PGRS"
/note="RV2741, (MTV002.06), len: 525 aa; Member of M.
tuberculosis PE_PGRS subfamily, similar to many eg.

Query Match 98.8%; Score 1213.8; DB 1; Length 56414;
Best Local Similarity 99.7%; Pred. No. 1.2e-165;
Matches 1226; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 ATCTTGACATTAATGAACTTTCTTACACCTGACAGCTATCGAGAGGGTAAATC 60
DB 38201 ATCTTGACATTAATGAACTTTCTTACACCTGACAGCTATCGAGAGGGTAAATC 38260
QY 61 ATGCGGTGGGTATTCGACGAGAACCAAAACAAAGCAATTCGGGTGGCCATCAACCCCG 120
DB 38261 ATGCGGTGGGTATTCGACGAGAACCAAAACAAAGCAATTCGGGTGGCCATCAACCCCG 38320
QY 121 GCGGCGTGGCGAACTACCCGTGTGCGCATGAGTGTCT ATCCAGGAGGTGGCCGA 180
DB 38321 GCGGCGTGGCGAACTACCCGTGTGCGCATGAGTGTCTATCCAGGAGGTGGCCGA 38380
QY 181 GAGGCGTGGCTATCATCGACGCGGATTTCAAAGGCGGCGCACTGCTGGCCACC 240
DB 38381 GAGGCGTGGCTATCATCGACGCGGATTTCAAAGGCGGCGCACTGCTGGCCACC 38440
QY 241 GCGGACAGGTGGGGCGGACGCTGATTTATGCTCAAGGTCAAAAGACGATAGCGGGG 300
DB 38441 GCGGACAGGTGGGGCGGACGCTGATTTATGCTCAAGGTCAAAAGACGATAGCGGGG 38500
QY 301 GAATAGCGCGCTGCGACGACGAGGCG-GATCTTGTTCAGTCTTGTGATTTGGCGCGCTCA 359
DB 38501 GAATAGCGCGCTGCGACGACGAGGCGAGTCTTGTGATTTGGCGCGCTCA 38560
QY 360 CGTGTGTGACACGAGTGTGTGATTCGCGACACGTCATTTCCCTACGAGACCGTTC 419
DB 38561 CGTGTGTGACACGAGTGTGTGATTCGCGACACGTCATTTCCCTACGAGACCGTTC 38620
QY 420 CAGACCGCGGAGGCGCATACCCGCTTGGCCGATGAGGGAATGCGCGGTGCACTC 479
DB 38621 CAGACCGCGGAGGCGCATACCCGCTTGGCCGATGAGGGAATGCGCGGTGCACTC 38680
QY 480 GCGGCCAGGTGGCGCTTACCACTGATGCAACCAAGGGGCGCGGTGTGTGATG 539
DB 38681 GCGGCCAGGTGGCGCTTACCACTGATGCAACCAAGGGGCGCGGTGTGTGATG 38740
QY 540 GCGGCGGTGCGCGCTGCAACCGCGGACGTCGTGTGATGCGCGCGGACCGCGGCG 599
DB 38741 GCGGCGGTGCGCGCTGCAACCGCGGACGTCGTGTGATGCGCGCGGACCGCGGCG 38800
QY 600 TACAAGGAGCCCGCATTCGCAACGCGCATGGGCGGACCGTTAGGCTTACGATCAAC 659
DB 38801 TACAAGGAGCCCGCATTCGCAACGCGCATGGGCGGACCGTTAGGCTTACGATCAAC 38860
QY 660 ATGACAAACTTGGCACTGCAAGCGGATTCGCGCGGATTCACATCGCTACTGA 719
DB 38861 ATGACAAACTTGGCACTGCAAGCGGATTCGCGCGGATTCACATCGCTACTGA 38920
QY 720 TCGGCTTACGAGTTCGAGGTGCGCTCAAAAGTGGCGGACTGTGATTTGGGCGCTGCTG 779
DB 38921 TCGGCTTACGAGTTCGAGGTGCGCTCAAAAGTGGCGGACTGTGATTTGGGCGCTGCTG 38980
QY 780 GTCGCGAGGCGCAAGGACCAAAATAGTCTGAATTCATTTCCGCAATGAACA 839
DB 38981 GTCGCGAGGCGCAAGGACCAAAATAGTCTGAATTCATTTCCGCAATGAACA 39040
QY 840 GGTGCGGTACTGTGTGATATGACCATGACGAGGCGGCTGTTTCAAGGCTCAGACCG 899
DB 39041 GGTGCGGTACTGTGTGATATGACCATGACGAGGCGGCTGTTTCAAGGCTCAGACCG 39100

QY 900 ACCACTACGACACCCGACGCTTGGCTGCAACGACACGCTTTTACTGCGGTGGCGAAC 959
DB 39101 ACCACTACGACACCCGACGCTTGGCTGCAACGACACGCTTTTACTGCGGTGGCGAAC 39160
QY 960 ATGCGCGCTGCGTGGCGAGACGTCGACCTACGGGTGACCAACGAGCATGCCGTAT 1019
DB 39161 ATGCGCGCTGCGTGGCGAGACGTCGACCTACGGGTGACCAACGAGCATGCCGTAT 39220
QY 1020 GTGCTGAGCTTGGCCGACCATGGCTGGCGGCGGCGTGGCGGATTCGCGACTAGCC 1079
DB 39221 GTGCTGAGCTTGGCCGACCATGGCTGGCGGCGGCGTGGCGGATTCGCGACTAGCC 39280
QY 1080 AAAGTCTTTGACACGACGAGAGGCGCTTACTGTCCGAAGGAGGTGGCCACGACTGGG 1139
DB 39281 AAAGTCTTTGACACGACGAGAGGCGCTTACTGTCCGAAGGAGGTGGCCACGACTGGG 39340
QY 1140 GTGCGCTTACCGACCGCGCGCGGCTGCTGGCGCTGACCTGCGCGCTGTTAGCGCGAG 1199
DB 39341 GTGCGCTTACCGACCGCGCGCGGCTGCTGGCGCTGACCTGCGCGCTGTTAGCGCGAG 39400
QY 1200 CACACGTCNGAGTAAGGAAACGATGATG 1229
DB 39401 CACACGTCNGAGTAAGGAAACGATGATG 39430

RESULT 12

LOCUS AB7606 1235 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 3 from Patent W09836089.
ACCESSION AB7606
VERSION AB7606.1 GI:6736246
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 1235)
AUTHORS
TITLE
JOURNAL
FLOHE LEPOLD (DE) SINGH MAHAVIR (DE)
FEATURES
Location/Qualifiers
source 1..1235
/organism="unidentified"
/db_xref="taxon:3264"
BASE COUNT 236 a 394 c 382 g 220 t 3 others
ORIGIN

Query Match 98.7%; Score 1213.4; DB 5; Length 1235;
Best Local Similarity 99.6%; Pred. No. 3.4e-165;
Matches 1225; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1 ATCTTGACATTAATGAACTTTCTTACACCTGACAGCTATCGAGAGGGTAAATC 60
DB 1 ATCTTGACATTAATGAACTTTCTTACACCTGACAGCTATCGAGAGGGTAAATC 60
QY 61 ATGCGGTGGGTATTCGACGAGAACCAAAACAAAGCAATTCGGGTGGCCATCAACCCCG 120
DB 61 ATGCGGTGGGTATTCGACGAGAACCAAAACAAAGCAATTCGGGTGGCCATCAACCCCG 120
QY 121 GCGGCGTGGCGGAACTACCCGTGTGCGCATGAGTGTCTATCCAGGAGGTGGCCGA 180
DB 121 GCGGCGTGGCGGAACTACCCGTGTGCGCATGAGTGTCTATCCAGGAGGTGGCCGA 180
QY 181 GAGGCTTGGCTATACGACGCGGATTTCAAAGGCGGACGCGCAACTGTGCGGAC 240
DB 181 GAGGCTTGGCTATACGACGCGGATTTCAAAGGCGGACGCGCAACTGTGCGGAC 240
QY 241 GCGGACAGGTGTGGGCGGACGCTGATTTATGCTCAAGGTCAAAAGACGATAGCGGCG 300
DB 241 GCGGACAGGTGTGGGCGGACGCTGATTTATGCTCAAGGTCAAAAGACGATAGCGGCG 300

```

OY 301 GAATACGGCCGCTGCGACAGGGG-GATCTGTTACGCTTGTGCAATTTGGCCGCTCA 359
DB 301 GATACGGCCGCTGCGACAGGGG-GATCTGTTACGCTTGTGCAATTTGGCCGCTCA 360
OY 360 CCGTGTGACCGATGCGCTTGTGATTCGGGACACGCTCAATTTGCCATGACAGCCGTG 419
DB 361 CCGTGTGACCGATGCGCTTGTGATTCGGGACACGCTCAATTTGCCATGACAGCCGTG 420
OY 420 CAGACCGCCGAAAGGCGCATACCCCTGCTTGGCCGCGATAGAGGAAGTCCCGGTGACATC 479
DB 421 CAGACCGCCGAAAGGCGCATACCCCTGCTTGGCCGCGATAGAGGAAGTCCCGGTGACATC 480
OY 480 GCGGCCGAGGTTGGGCGCTTACCACTGATGCAACCAAGGGGCGCGGTGCTGATG 539
DB 481 GCGGCCGAGGTTGGGCGCTTACCACTGATGCAACCAAGGGGCGCGGTGCTGATG 540
OY 540 GCGGGGTCGCGGCGCTGCAACCGGCGAGCTGTGTGATGCGGCGCGGACCGCCGCG 599
DB 541 GCGGGGTCGCGGCGCTGCAACCGGCGAGCTGTGTGATGCGGCGCGGACCGCCGCG 600
OY 600 TACAAAGCAGCCCGCATTCGCCAAAGGCGGCGGCGGCGGCTTACAGCTTACACATCAAC 659
DB 601 TACAAAGCAGCCCGCATTCGCCAAAGGCGGCGGCGGCGGCTTACAGCTTACACATCAAC 660
OY 660 ATGCAAACTTCGCGACTGCGACGCGGAGTTCTGCGGCGGCTATCCACACTGCTACTCA 719
DB 661 ATGCAAACTTCGCGACTGCGACGCGGAGTTCTGCGGCGGCTATCCACACTGCTACTCA 720
OY 720 TCGGCTTACGAGCTGAGGGTGGCTTCAAAAGTGGCGGCGGCTGATTTGGGCGCTCTG 779
DB 721 TCGGCTTACGAGCTGAGGGTGGCTTCAAAAGTGGCGGCGGCTGATTTGGGCGCTCTG 780
OY 780 GTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 839
DB 781 GTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 840
OY 840 GGTGCGGCTAGTGTGATATAGCATGACCAAGGCGGCTGTTTGAAGGCTCAGACCG 899
DB 841 GGTGCGGCTAGTGTGATATAGCATGACCAAGGCGGCTGTTTGAAGGCTCAGACCG 900
OY 900 AACACTAGACACCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 959
DB 901 AACACTAGACACCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
OY 960 ATGCGCGCTGCGGTCGCAAGAGCTGACCTACGCGCTGACCAACGCGAGATGCGGTAT 1019
DB 961 ATGCGCGCTGCGGTCGCAAGAGCTGACCTACGCGCTGACCAACGCGAGATGCGGTAT 1020
OY 1020 GTGCTGAGCTTCCGACCATGCTGCGGCGGCGGCGGCTGCGGCTGCAATCGGCACTAGCC 1079
DB 1021 GTGCTGAGCTTCCGACCATGCTGCGGCGGCGGCGGCTGCGGCTGCAATCGGCACTAGCC 1080
OY 1080 AAGGCTTTTCAGCGTACGAAGGAGGCGGCTTACCTGCGAAGGGGTGGCGACCGACTGGGG 1139
DB 1081 AAGGCTTTTCAGCGTACGAAGGAGGCGGCTTACCTGCGAAGGGGTGGCGACCGACTGGGG 1140
OY 1140 GTGCGCTTACGCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1199
DB 1141 GTGCGCTTACGCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1200
OY 1200 CACAGCTCAGAGTAAGGAGGAGGATGATG 1229
DB 1201 CACAGCTCAGAGTAAGGAGGAGGATGATG 1230

```

RESULT 13

```

LOCUS A89746 1235 bp DNA
DEFINITION Sequence 3 from Patent WO9832862.
ACCESSION A89746
VERSION A89746.1 GI:6738280
KEYWORDS
SOURCE unidentified.

```

```

ORGANISM unidentified
REFERENCE 1 (bases 1 to 1235)
AUTHORS Flohe, L. and Singh, M.
TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
JOURNAL Patent: WO 9832862-A-30-JUL-1998;
FLOHE LEOBOLD (DE); SINGH MAHAVIR (DE)
FEATURES
    source
        location/Qualifiers
            1..1235
                /organism="unidentified"
                /db_xref="taxon:32644"
BASE COUNT 236 a 394 c 382 g 220 t 3 others
ORIGIN
Query Match 98.7% Score 1213.4; DB 5; Length 1235;
Best Local Similarity 99.6% Pred. No. 3.4e-165;
Matches 1225; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
OY 1 ATCTGCAAGATTATCGAAGTTCTTCACTGAAAGCTACAGTATGAGAGGGGTATATC 60
DB 1 ATCTGCAAGATTATCGAAGTTCTTCACTGAAAGCTACAGTATGAGAGGGGTATATC 60
OY 61 ATGCGCGTGGTATTCGACGAGACCAAAAAACAAGATTCCGGGTGCCATACCCCG 120
DB 61 ATGCGCGTGGTATTCGACGAGACCAAAAAACAAGATTCCGGGTGCCATACCCCG 120
OY 121 GCGGCGCTCGCGGAACTAACCCGCTGTGCGCATGAGTGTCTATCCAGGAGGTGCCGA 180
DB 121 GCGGCGCTCGCGGAACTAACCCGCTGTGCGCATGAGTGTCTATCCAGGAGGTGCCGA 180
OY 181 GAGGCGCTCGCTATCACGACGCGGATTTCAAGGCGGAGCGGCGCACTGGTGGGCGAC 240
DB 181 GAGGCGCTCGCTATCACGACGCGGATTTCAAGGCGGAGCGGCGCACTGGTGGGCGAC 240
OY 241 GCGGACAGGTGTGGGCGGAGCGTGTATTTCTCAAGTCAAAAGAACGATAGCGCG 300
DB 241 GCGGACAGGTGTGGGCGGAGCGTGTATTTCTCAAGTCAAAAGAACGATAGCGCG 300
OY 301 GAATACGGCCGCTGCGACAGGGG-GATCTGTTACGCTTGTGCAATTTGGCCGCTCA 359
DB 301 GAATACGGCCGCTGCGACAGGGG-GATCTGTTACGCTTGTGCAATTTGGCCGCTCA 360
OY 360 CCGTGTGACCGATGCGCTTGTGATTCGGGACACGCTCAATTTGCCATGACAGCCGTG 419
DB 361 CCGTGTGACCGATGCGCTTGTGATTCGGGACACGCTCAATTTGCCATGACAGCCGTG 420
OY 420 CAGACCGCCGAAAGGCGCATACCCCTGCTTGGCCGCGATAGAGGAAGTCCCGGTGACATC 479
DB 421 CAGACCGCCGAAAGGCGCATACCCCTGCTTGGCCGCGATAGAGGAAGTCCCGGTGACATC 480
OY 480 GCGGCCGAGGTTGGGCGCTTACCACTGATGCAACCAAGGGGCGCGGTGCTGATG 539
DB 481 GCGGCCGAGGTTGGGCGCTTACCACTGATGCAACCAAGGGGCGCGGTGCTGATG 540
OY 540 GCGGGGTCGCGGCGCTGCAACCGGCGAGCTGTGTGATGCGGCGCGGACCGCCGCG 599
DB 541 GCGGGGTCGCGGCGCTGCAACCGGCGAGCTGTGTGATGCGGCGCGGACCGCCGCG 600
OY 600 TACAAAGCAGCCCGCATTCGCCAAAGGCGGCGGCGGCGGCTTACAGCTTACACATCAAC 659
DB 601 TACAAAGCAGCCCGCATTCGCCAAAGGCGGCGGCGGCGGCTTACAGCTTACACATCAAC 660
OY 660 ATGCAAACTTCGCGACTGCGACGCGGAGTTCTGCGGCGGCTATCCACACTGCTACTCA 719
DB 661 ATGCAAACTTCGCGACTGCGACGCGGAGTTCTGCGGCGGCTATCCACACTGCTACTCA 720
OY 720 TCGGCTTACGAGCTGAGGGTGGCTTCAAAAGTGGCGGCGGCTGATTTGGGCGCTCTG 779
DB 721 TCGGCTTACGAGCTGAGGGTGGCTTCAAAAGTGGCGGCGGCTGATTTGGGCGCTCTG 780
OY 780 GTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 839

```

```

Db 781 GTCCAGAGCCCAAGGACCCAAATTAGTCTCAATTCACTTGTGCGCATATGAACCA 840
QY 840 GTGCGGTACTGTGATATATAGCCATGACACGAGGCGGTGTTTGAAGGCTCAGCAGC 899
Db 841 GTGCGGTACTGTGATATATAGCCATGACACGAGGCGGTGTTTGAAGGCTCAGCAGC 900
QY 900 ACCACCTACGACACCCGAGCTTCCGCTGACAGACAGCTGTTTACTGCTGGGCAAC 959
Db 901 ACCACCTACGACACCCGAGCTTCCGCTGACAGACAGCTGTTTACTGCTGGGCAAC 960
QY 960 ATGCGCGCTCGGTGCGGAGAGCTGACCTAGCCGTGACAAAGGCGAGATGCCGTAT 1019
Db 961 ATGCGCGCTCGGTGCGGAGAGCTGACCTAGCCGTGACAAAGGCGAGATGCCGTAT 1020
QY 1020 GTGCTGAGCTTCCGACCATGCTGCGGGGCGGCGCTGCGGCTGGAATCCGGCAATAGCC 1079
Db 1021 GTGCTGAGCTTCCGACCATGCTGCGGGGCGGCGCTGCGGCTGGAATCCGGCAATAGCC 1080
QY 1080 AAGGCTTTTCAGACGACGAAAGGGGCTTACTGTCGGAAGGGGTGGCCACCGACTGGGG 1139
Db 1081 AAGGCTTTTCAGACGACGAAAGGGGCTTACTGTCGGAAGGGGTGGCCACCGACTGGGG 1140
QY 1140 GTGCGGTTCACGAGCGCCGCGCAGCGTCTGAGCTCTCAGCGGCTGTTAGCGCGAG 1199
Db 1141 GTGCGGTTCACGAGCGCCGCGCAGCGTCTGAGCTCTCAGCGGCTGTTAGCGCGAG 1200
QY 1200 CACACGTGNGAGTAAGGAAGCGATGATG 1229
Db 1201 CACACGTGNGAGTAAGGAAGCGATGATG 1230

RESULT 14
LOCUS A87608 1228 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 5 from Patent WO9836089.
ACCESSION A87608
VERSION A87608.1 GI:6736248
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1. (bases 1 to 1228)
AUTHORS Flohe, L. and Singh, M.
TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL Patent: WO 9836089-A 20-AUG-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
source 1..1228
location/Organism
/organism="unidentified"
/ds_xref="taxon:32644"
BASE COUNT 236 a 391 c 382 g 219 t
ORIGIN

```

```

Query Match 98.6%; Score 1211.8; DB 5; Length 1228;
Best Local Similarity 99.7%; Pired. No. 5; 7e-165;
Matches 1224; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```

```

QY 241 GCCGACAGGTGTGGCGCCGACGCTGATTTATGCTCAAGGTCAAAAGAACGATAGCGCG 300
Db 241 GCCGACAGGTGTGGCGCCGACGCTGATTTATGCTCAAGGTCAAAAGAACGATAGCGCG 300
QY 301 GAATACGGCGCGCTGGACACGGGC-GATCTTGTACAGTTCTTTCGATTTGGCGCGTCA 359
Db 301 GAATACGGCGCGCTGGACACGGGCAGATCTTGTTCACGTTCTTTCGATTTGGCGCGTCA 360
QY 360 CGTGTTCACGATCGCTTGTGATTCGGGACACGCAATTCGCTACGAGACCGTC 419
Db 361 CGTGTTCACGATCGCTTGTGATTCGGGACACGCAATTCGCTACGAGACCGTC 420
QY 420 CAGACCGCGGAAAGCGACATACCCCTGCTTGCCTGATGAGCGAAGTCCGCGTCACTC 479
Db 421 CAGACCGCGGACCGACATACCCCTGCTTGCCTGATGAGCGAAGTCCGCGTCACTC 480
QY 480 GCGCGCCAGGTTGGCGCTTACCACTGATGCGAACCCAAAGGGGCGCGGTGTCTGATG 539
Db 481 GCGCGCCAGGTTGGCGCTTACCACTGATGCGAACCCAAAGGGGCGCGGTGTCTGATG 540
QY 540 GCGCGGCTGCGCGCGCTGCAACCGCGCGAGTGTGTGATGCGCGCGCGACCGCGCG 599
Db 541 GCGCGGCTGCGCGCGCTGCAACCGCGCGAGTGTGTGATGCGCGCGCGACCGCGCG 600
QY 600 TACACGCAAGCCCGCATCGCAAGGCGCATGGCGCGACCGTTACGGTTTACATCAAC 659
Db 601 TACACGCAAGCCCGCATCGCAAGGCGCATGGCGCGACCGTTTACATCAAC 660
QY 660 ATGACAAACTTGGCACTCGACGCGGAGTTGTGGCGCGGATCCACACTGGCTACTCA 719
Db 661 ATGACAAACTTGGCACTCGACGCGGAGTTGTGGCGCGGATCCACACTGGCTACTCA 720
QY 720 TCGGCTACGAGCTCGAGGCTGCGTCAACGCGCGACCTGGTGTGTTGGGCGCGCTG 779
Db 721 TCGGCTACGAGCTCGAGGCTGCGTCAACGCGCGACCTGGTGTGTTGGGCGCGCTG 780
QY 780 GTGCGAGGCGCCAGGACCCCAATTAAGTCTGCAATTAAGTCTGCGCATATGAACCA 839
Db 781 GTGCGAGGCGCCAGGACCCCAATTAAGTCTGCAATTAAGTCTGCGCATATGAACCA 840
QY 840 GTGCGGTACTGTGATATATAGCCATGACACGAGGCGGCTGTTTGAAGGCTCAGACCG 899
Db 841 GTGCGGTACTGTGATATATAGCCATGACACGAGGCGGCTGTTTGAAGGCTCAGACCG 900
QY 900 ACCACCTACGACACCCGAGCTTCCGCTGACAGACAGCTGTTTACTGCGGGGCAAC 959
Db 901 ACCACCTACGACACCCGAGCTTCCGCTGACAGACAGCTGTTTACTGCGGGGCAAC 960
QY 960 ATGCGCGCTCGGTGCGGAGAGCTGACCTAGCGCTGACCAACGCGAGATGCCGTAT 1019
Db 961 ATGCGCGCTCGGTGCGGAGAGCTGACCTAGCGCTGACCAACGCGAGATGCCGTAT 1020
QY 1020 GTGCTGAGCTTCCGACCATGCTGCGGGGCGGCGCTGCGGCTGGAATCCGGCAATAGCC 1079
Db 1021 GTGCTGAGCTTCCGACCATGCTGCGGGGCGGCGCTGCGGCTGGAATCCGGCAATAGCC 1080
QY 1080 AAGGCTTTTCAGACGACGAAAGGGGCTTACTGTCGGAAGGGGTGGCCACCGACTGGGG 1139
Db 1081 AAGGCTTTTCAGACGACGAAAGGGGCTTACTGTCGGAAGGGGTGGCCACCGACTGGGG 1140
QY 1140 GTGCGGTTCACGAGCGCCGCGCAGCGTCTGAGCTCTCAGCGGCTGTTAGCGCGAG 1199
Db 1141 GTGCGGTTCACGAGCGCCGCGCAGCGTCTGAGCTCTCAGCGGCTGTTAGCGCGAG 1200
QY 1200 CACACGTGNGAGTAAGGAAGCGATGA 1227
Db 1201 CACACGTGNGAGTAAGGAAGCGATGA 1228

RESULT 15
A89748 1228 bp DNA PAT 22-JAN-2000
LOCUS

```



```

DEFINITION   Sequence 5 from Patent WO9832862.
ACCESSION    A89748
VERSION      A89748.1 GI:6738282
KEYWORDS     unidentified.
SOURCE       unidentified.
ORGANISM     unidentified.
REFERENCE    1 (bases 1 to 1228)
AUTHORS     Flohe, L. and Singh, M.
TITLE       L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
JOURNAL      Patent: WO 9832862-A 30-JUL-1998;
            FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
    source
        1..1228
            /organism="unidentified"
            /db_xref="taxon:32644"
BASE COUNT   236 a 391 c 382 g 219 t
ORIGIN
Query Match      98.6%; Score 1211.8; DB 5; Length 1228;
Best Local Similarity 99.7%; Pred. No. 5.7e-165;
Matches 1224; Conservative 0; Mismatches 4; Indels 1; Gap: 1;
QY 1 ATCTTGAGATTATGCAACTTTCTTACACTGAGCGTACATATCGAAGGGTATC 60
DB 1 ATCTTGAGATTATGCAACTTTCTTACACTGAGCGTACATATCGAAGGGTATC 60
QY 61 ATGCGCGTGGTATTCGACGCGAGACCAAAACAGCAATTCGGGGTGGCATCACCCG 120
DB 61 ATGCGCGTGGTATTCGACGCGAGACCAAAACAGCAATTCGGGGTGGCATCACCCG 120
QY 121 GCCGGCGTCGGAACTAACCCGTCGTGGCCATGAGGTGCTCATCCAGGAGTGGCGGA 180
DB 121 GCCGGCGTCGGAACTAACCCGTCGTGGCCATGAGGTGCTCATCCAGGAGTGGCGGA 180
QY 181 GAGGGCTCGGCTATACCGGAGCGGGATTTCAAGGCGGACCGCCGCACTGGTCGCACC 240
DB 181 GAGGGCTCGGCTATACCGGAGCGGGATTTCAAGGCGGACCGCCGCACTGGTCGCACC 240
QY 241 GCCGACGAGGTGTGGGCGGACGCTGATTTATGCTCAAGGTCAAGAACCGATAGCGGC 300
DB 241 GCCGACGAGGTGTGGGCGGACGCTGATTTATGCTCAAGGTCAAGAACCGATAGCGGC 300
QY 301 GAATPAGGCGCGCTGCGACACGGGC- GATCTGTTCACGTTCTTGCAATTTGGCGCGTCA 359
DB 301 GAATPAGGCGCGCTGCGACACGGGC- GATCTGTTCACGTTCTTGCAATTTGGCGCGTCA 359
QY 360 CGATGCTGACGCGATGCTGTGGATTCGGGACACGTCATGCTACGAGACCGTTC 419
DB 360 CGATGCTGACGCGATGCTGTGGATTCGGGACACGTCATGCTACGAGACCGTTC 419
QY 420 CAGACGCGCCGAGGCGGACGCTGCTTGGCCCGATGAGCGAAGTCGCGGTCGACTC 479
DB 420 CAGACGCGCCGAGGCGGACGCTGCTTGGCCCGATGAGCGAAGTCGCGGTCGACTC 479
QY 480 GCCGCCAGGTGGCGTTACCACTGATGCGAACCCAGGGGGCGCGGTGTGTGATG 539
DB 480 GCCGCCAGGTGGCGTTACCACTGATGCGAACCCAGGGGGCGCGGTGTGTGATG 539
QY 540 GGGGGGGTGGCGGCGTGAACCGGGCGAGTGTGTGATTCGGCCCGGACACCGCGGC 599
DB 540 GGGGGGGTGGCGGCGTGAACCGGGCGAGTGTGTGATTCGGCCCGGACACCGCGGC 599
QY 541 GGGGGGGTGGCGGCGTGAACCGGGCGAGTGTGTGATTCGGCCCGGACACCGCGGC 600
DB 541 GGGGGGGTGGCGGCGTGAACCGGGCGAGTGTGTGATTCGGCCCGGACACCGCGGC 600
QY 600 TACAAGCGACCGCGCATGCGCAAGGCGATGGCGGACCGTACGGTTAGACATCAAC 659
DB 600 TACAAGCGACCGCGCATGCGCAAGGCGATGGCGGACCGTACGGTTAGACATCAAC 659
QY 660 ATGACAACTTGGCACTGACGCGGAGTGTGCGCGGATCCACACTCGCTACTCA 719
DB 660 ATGACAACTTGGCACTGACGCGGAGTGTGCGCGGATCCACACTCGCTACTCA 719
QY 720 TGGGCTAGAGCTCGAGGGTGGCGTCAACGTCGCGACTGTGTGATTTGGGGCGCTCTG 779
DB 720 TGGGCTAGAGCTCGAGGGTGGCGTCAACGTCGCGACTGTGTGATTTGGGGCGCTCTG 779

```

```

DB 721 TGGGCTAGAGCTCGAGGGTGGCGTCAACGTCGCGACTGTGTGATTTGGGGCGCTCTG 780
QY 780 GTGCCAGGCGCCCAAGGACACCAATAGTCTGGAATTCATTTGGCGCATATGAACCA 839
DB 781 GTGCCAGGCGCCCAAGGACACCAATAGTCTGGAATTCATTTGGCGCATATGAACCA 840
QY 840 GTGCCAGTCTGTGATATGATGATGACGACGAGGGGGGTGTTGGAAGGCTCAGACCG 899
DB 841 GTGCCAGTCTGTGATATGATGATGACGACGAGGGGGGTGTTGGAAGGCTCAGACCG 900
QY 900 ACCACCTAGACACCGACCGAGCTTGGCGGTGCAAGACAGCTTTTACTGCTGGCGAAC 959
DB 901 ACCACCTAGACACCGACCGAGCTTGGCGGTGCAAGACAGCTTTTACTGCTGGCGAAC 960
QY 960 ATGCCCGCTCGTGGCGGAGAGCTGACCTACGCGCTGACCAAGCGCATGCTGCTAT 1019
DB 961 ATGCCCGCTCGTGGCGGAGAGCTGACCTACGCGCTGACCAAGCGCATGCTGCTAT 1020
QY 1020 GTGCTGAGCTTGGCGGACCATGATGATGCGGGGCGGTGCGGTCGATCCGACTAGCC 1079
DB 1021 GTGCTGAGCTTGGCGGACCATGATGATGCGGGGCGGTGCGGTCGATCCGACTAGCC 1080
QY 1080 AAGGCTTTTCAACGACGAGAGGGCGTTACTGTCCGAACGGGTGGCCACCGACTGGGG 1139
DB 1081 AAGGCTTTTCAACGACGAGAGGGCGTTACTGTCCGAACGGGTGGCCACCGACTGGGG 1140
QY 1140 GTGCGTTACGAGAGCGCGCGGCGGCTGCTGACTGCTGCGCGCTGCTGCTGCTGCTG 1199
DB 1141 GTGCGTTACGAGAGCGCGCGGCGGCTGCTGACTGCTGCGCGCTGCTGCTGCTGCTG 1200
QY 1200 CACAGCTGAGAGTAAGGAGACGATGA 1227
DB 1201 CACAGCT GAGAGTAAGGAGACGATGA 1228

```

Search completed: June 22, 2000, 14:55:19
Job time: 17734 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 15:09:11 ; Search time 458.59 Seconds

(without alignments)
670.503 Million cell updates/sec

Title: US-09-362-485-7

Perfect score: 1229
Sequence: 1 ATCTTGCAATTATCGAAC.....GAGTAAGGAGAGCATGATG 1229

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 08

Listing first 45 summaries

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1199.4	97.6	1245	1 V49626	Mycobacterium tube
2	1199.4	97.6	1260	1 V49510	Mycobacterium sp.
3	1199.4	97.6	1260	1 V49625	Mycobacterium tube
4	454	36.9	682	1 V49511	Mycobacterium mar
5	265.2	21.6	1125	1 N91423	Sequence of heat-r
6	203.6	16.6	28171	1 V52155	Streptococcus pneu
7	166.4	13.5	9280	1 V74442	Staphylococcus aur
8	89.2	7.3	1074	1 T67971	H. pylori membrane
9	69.2	5.6	544	1 V43039	Streptococcus pneu
10	56.8	4.6	31096	1 V74370	Streptococcus aur
11	49.2	4.0	390	1 Q21833	Randomising oligon
12	49.2	4.0	390	1 Q36859	PCR primer for 5'
13	48.8	4.0	15872	1 T68715	Streptomyces venez
14	46.4	3.8	985	1 V44439	Mycobacterium tube
15	46.4	3.8	985	1 V64548	M. tuberculosis im
16	45.4	3.7	833	1 Q64203	snab gene encoding
17	45.4	3.7	5392	1 Q64201	Sequence comprisin
18	45	3.7	3946	1 V93610	Mycobacterium tube
19	44.8	3.6	535	1 V44428	Mycobacterium tube
20	44.8	3.6	535	1 V64537	M. tuberculosis im
21	44.4	3.6	2582	1 T73117	Actinoplanes sp. a
22	44.4	3.6	17955	1 V56642	Actinoplanes sp. a
23	43.6	3.5	28598	1 T06769	Sorangium cellulos
24	43.6	3.5	28598	1 T89956	Sorangium cellulos
25	43.6	3.5	49377	1 V05287	The soraphen biosy
26	42.8	3.5	114955	1 X53491	Human adenovirus A1
27	42.4	3.4	1833	1 Q64206	snbr gene encoding
28	42.4	3.4	2185	1 V84066	Clone p5-4 encodin
29	42.4	3.4	2185	1 Q09365	S. lividans protea
30	42.4	3.4	24379	1 T93095	Streptomyces freno
31	42.4	3.4	24379	1 V25925	Streptomyces roseo
32	42.2	3.4	329	1 V44425	Mycobacterium tube
33	42.2	3.4	329	1 V64534	M. tuberculosis im
34	41.8	3.4	882	1 V44403	Mycobacterium tube

35	41.8	3.4	882	1 V64512	M. tuberculosis im
36	41	3.3	3946	1 T93610	Mycobacterium tube
37	41	3.3	29879	1 Q46806	erya region of S.
38	40.8	3.3	2151	1 Q76252	HSV-2 protease, IC
39	40.8	3.3	2151	1 Q76261	HSV-2 protease/ICP
40	40.8	3.3	2472	1 Q84671	HSV-2 UL26 gene. N
41	40.8	3.3	20387	1 V62159	HSV-2 strain SB5 C
42	40.8	3.3	26338	1 V62134	HSV-2 strain SB5 C
43	40.8	3.3	117213	1 V62176	HSV-2 strain SB5 C
44	40.6	3.3	4257	1 V10362	Infected cell prot
45	40.6	3.3	4257	1 V68520	The nucleotide seq

ALIGNMENTS

RESULT	1
V49626	V49626 standard; DNA; 1245 BP.
AC	V49626;
DT	20-NOV-1998 (first entry)
DE	Mycobacterium tuberculosis L-Alanine dehydrogenase 40 KD antigen.
KW	ss; Alanine dehydrogenase; tuberculosis; TB; mycobacteria; L-alanine.
OS	Mycobacterium tuberculosis.
PN	MO9836089-A2.
PD	20-AUG-1998.
PF	29-JAN-1998: E00483.
PR	29-JAN-1997: BP-101338.
PA	(FLOH/) FLOHE L.
PI	Flohe L. Hutter B. Kolk A. Singh M.
DR	WPI: 98-457123/39.
PT	Diagnosis of, e.g. tuberculosis from alanine dehydrogenase activity
PT	- useful for, e.g. for diagnosis, differentiation of strains,
PT	monitoring vaccination and identification of mycobacterial
PT	inhibitors
PS	Disclosure; Fig 3.19; 55pp; German.
CC	The Mycobacterium tuberculosis L-alanine dehydrogenase (AlaDH) is used in
CC	the production of kits for diagnosing tuberculosis (TB) and other
CC	mycobacterial infections in humans or animals. Kits are used for direct
CC	diagnosis of TB on clinical samples (e.g. body fluids) and can
CC	differentiate between pathogenic and non-virulent strains, e.g. for
CC	identifying pathogens of the M. tuberculosis (M. t.) complex. The kit may
CC	also be used to identify substances that inhibit mycobacteria, for
CC	combating epidemics and for vaccination follow-up. Oligonucleotides
CC	derived from AlaDH are used similarly in diagnostic hybridisation tests,
CC	also for culture confirmation of isolated strains and for chromosome
CC	fingerprinting to detect/differentiate between mycobacteria, and for
CC	L-alanine-specific biotransformation reactions. AlaDH is an early
CC	antigen, present extracellularly after only a few days of growth, making
CC	it an ideal drug target.
SO	Sequence 1245 BP; 238 A; 398 C; 387 G; 222 T;
Query Match	97.6%; Score 1199.4; DB 1; Length 1245;
Best Local Similarity	99.3%; Pred. No. 1.5e+240;
Matches 1227; Conservative	0; Mismatches 2; Indels 7; Gaps 2;
QY	1 ATCTTGCAATTATCGAAGTTCTTCACTGACGTAACGTAATCGAGGCGGTAATC 60
DB	1 ATCTTGCAATTATCGAAGTTCTTCACTGACGTAACGTAATCGAGGCGGTAATC 60
QY	61 ATCTTGCAATTATCGAAGTTCTTCACTGACGTAACGTAATCGAGGCGGTAATC 114
DB	61 ATCTTGCAATTATCGAAGTTCTTCACTGACGTAACGTAATCGAGGCGGTAATC 120
QY	115 ACCCGCGCGCGCGCGGTAACCGTGTGGCATGAGGTGCTCATCCAGGAGGT 174
DB	115 ACCCGCGCGCGCGCGGTAACCGTGTGGCATGAGGTGCTCATCCAGGAGGT 180
QY	175 GCGGAGAGGCGTTCGCTATCACCAGCGGAGTTTAAGGCGGCGGCAACTGGTC 234
DB	175 GCGGAGAGGCGTTCGCTATCACCAGCGGAGTTTAAGGCGGCGGCAACTGGTC 240

```

QY 235 GGCACCGCCGACAGGTGGGCGGACGCTGATTATGCTCAAGGCTCAAGACCGATA 294
DB 241 GGCACCGCCGACAGGTGGGCGGACGCTGATTATGCTCAAGGCTCAAGACCGATA 300
QY 295 GCGGCGGAATAGCGGCGCTGGGACACGGGCG-CATCTTGTCACTGTTTGAATTTGGCC 353
DB 301 GCGGCGGAATAGCGGCGCTGGGACACGGGCGAGATCTGTTCACTGTTTGAATTTGGCC 360
QY 354 GCGTCACTGCTTGCACCGCATGGTGTGATTCGGGACACCGCATTTGGCTACGAG 413
DB 361 GCGTCACTGCTTGCACCGCATGGTGTGATTCGGGACACCGCATTTGGCTACGAG 420
QY 414 ACCGTCAACACCGCCGGAAGGCGCACTACCTGCTGCCCCGATGAGCGAAGTCGCGGT 473
DB 421 ACCGTCAACACCGCCGGAAGGCGCACTACCTGCTGCCCCGATGAGCGAAGTCGCGGT 480
QY 474 CGACTGCGCGCCAGGTTGGCGCTTACCACTGATTCGGAACCCAGGGGCGCGGTGTG 533
DB 481 CGACTGCGCGCCAGGTTGGCGCTTACCACTGATTCGGAACCCAGGGGCGCGGTGTG 540
QY 534 CTGATGGGCGGGGTGCGGCGCTGCAACCGGCGGACGCTGCTGATTCGGGCGCGGAC 593
DB 541 CTGATGGGCGGGGTGCGGCGCTGCAACCGGCGGACGCTGCTGATTCGGGCGCGGAC 600
QY 594 GCGGCTACACGACGACCGCGCATGCGCAACGGCATGGGCGGCGCATGCGTTAGAC 653
DB 601 GCGGCTACACGACGACCGCGCATGCGCAACGGCATGGGCGGCGCATGCGTTAGAC 660
QY 654 ATCAACATGACAAATTTGGGCAACTGCAAGCGGCACTTCTGGGCGGGAATCCACTGCG 713
DB 661 ATCAACATGACAAATTTGGGCAACTGCAAGCGGCACTTCTGGGCGGGAATCCACTGCG 720
QY 714 TACTCATCGGCTTACGAGCTCGAGGCTGCGTCAACGCTGCGGACCTGCTGATTTGGGCG 773
DB 721 TACTCATCGGCTTACGAGCTCGAGGCTGCGTCAACGCTGCGGACCTGCTGATTTGGGCG 780
QY 774 GTCTGTGCTCCAGGCGCCCAAGGCAATTAATGCTGCAATTCCTGCTGCGGCAATG 833
DB 781 GTCTGTGCTCCAGGCGCCCAAGGCAATTAATGCTGCAATTCCTGCTGCGGCAATG 840
QY 834 AAACGAGTGGGCTTACTGGTGGATATAGCATGACGACGAGGCGGCTTTCGAAGCTCA 893
DB 841 AAACGAGTGGGCTTACTGGTGGATATAGCATGACGACGAGGCGGCTTTCGAAGCTCA 900
QY 894 CGACGACCACTTACGACACCGGAGCTTCGCGTGCACGACAGCGCTGTTTACTGCTG 953
DB 901 CGACGACCACTTACGACACCGGAGCTTCGCGTGCACGACAGCGCTGTTTACTGCTG 960
QY 954 GCGAATATGCGGCTCGGCTGCGGAGGAGAGTGCACCTACGCGCTGACCAACGCGAGATG 1013
DB 961 GCGAATATGCGGCTCGGCTGCGGAGGAGAGTGCACCTACGCGCTGACCAACGCGAGATG 1020
QY 1014 CCGTATGCTGAGAGTTCGCGACATGCTGCTGGGCGGCGGCTGCGAATCCGGCA 1073
DB 1021 CCGTATGCTGAGAGTTCGCGACATGCTGCTGGGCGGCGGCTGCGAATCCGGCA 1080
QY 1074 CTAGCCAAAGGCTTTCGACGACGAGAGGCGCTTACTGTCGGAACGGGTGGCCACGAC 1133
DB 1081 CTAGCCAAAGGCTTTCGACGACGAGAGGCGCTTACTGTCGGAACGGGTGGCCACGAC 1140
QY 1134 CTGGGGGTGCGCTTACGACGACGCGGCGGCTGCTGCTGCTGCTGCTGCTTAC 1193
DB 1141 CTGGGGGTGCGCTTACGACGACGCGGCGGCTGCTGCTGCTGCTGCTGCTTAC 1200
QY 1194 GCGGAGCACACGCTGNGAGTAAGGGAAGGATGATG 1229
DB 1201 GCGGAGCACACGCTGNGAGTAAGGGAAGGATGATG 1236

```

RESULT 2
 ID V49510 standard; DNA: 1260 BP.
 AC V49510;

DE 20-OCT-1998 (first entry)
 DE Mycobacterium sp. AlADH DNH.
 KM Alanine dehydrogenase; AlADH; ADH; diagnosis; tuberculosis; pathogen;
 KM swimmers disease; vaccine; epidemic; infection; identification; ss.
 OS Mycobacterium sp.
 PN M09832862-A2.
 PD 30-JUL-1998.
 PF 29-JAN-1998; E00484.
 PR 29-JAN-1997; EP-101339.
 PA (FLOH) FLOHE L.
 PA Flohe L, Hutter B, Kolk A, Singh M;
 DR WPI: 98-427958/36.
 PT Nucleic acid encoding alanine dehydrogenase of Mycobacterium marinum
 PT - used for diagnosis of tuberculosis and other mycobacterial
 PT diseases, also for treatment and prevention, for drug screening and
 PT for bio-transformation
 PS Disclosure: Page 11: 57P; German.
 CC This sequence encodes an alanine dehydrogenase (ADH) protein isolated
 CC from a Mycobacterium sp. This protein is used to diagnose tuberculosis
 CC and other mycobacterial infections (including 'swimmers' disease', caused
 CC by M. marinum, a fish pathogen) in humans or animals. The protein can
 CC also be used for control of epidemics and for vaccination, to screen for
 CC agents with anti-mycobacterial activity, and in bio-transformations that
 CC are specific for L-alanine. Also mycobacteria can be identified by
 CC analysis of genomic *fli* sequences. ADH is an antigen that is secreted
 CC early during infection
 SQ Sequence 1260 BP: 243 A: 403 C: 389 G: 225 T:

Query Match 97.6%; Score 1199.4; DB 1; Length 1260;
 Best Local Similarity 99.3%; Pred. No. 1.5e-240;
 Matches 1227; Conservative 0; Mismatches 2; Indels 7; Gaps 2;

```

QY 1 ATCTGCGATTAATCGAATCTTCTTCACTGAAAGCTATGATTCGAGAGGGTATTC 60
DB 16 ATCTGCGATTAATCGAATCTTCTTCACTGAAAGCTATGATTCGAGAGGGTATTC 75
QY 61 ATGCGGCTGCGATTCGACGACGACGCAAAACAAAG-----AATTCGGGTGGCCATC 114
DB 76 ATGCGGCTGCGATTCGACGACGACGCAAAACAAAGATTCGATTCGGGTGGCCATC 135
QY 115 ACCCGCGCGGCTGCGGCAATACCGGCTGCGGCAATGAGGTCGATCCAGGACGT 174
DB 136 ACCCGCGCGGCTGCGGCAATACCGGCTGCGGCAATGAGGTCGATCCAGGACGT 195
QY 175 GCGGAGAGGGCTGCGGCAATACCGGCTGCGGCAATGAGGTCGATCCAGGACGT 234
DB 196 GCGGAGAGGGCTGCGGCAATACCGGCTGCGGCAATGAGGTCGATCCAGGACGT 255
QY 235 GGCACCGCGGACGAGTGGGCGGACGCTGATTTATGCTCAAGGTAAGAAACCGATA 294
DB 256 GGCACCGCGGACGAGTGGGCGGACGCTGATTTATGCTCAAGGTAAGAAACCGATA 315
QY 295 GCGGCGGAATAGCGGCGCTGCGACACGGGCG-CATCTTGTCACTGTTTGAATTTGGCC 353
DB 316 GCGGCGGAATAGCGGCGCTGCGACACGGGCG-CATCTTGTCACTGTTTGAATTTGGCC 375
QY 354 GCGTCACTGCTTGCACCGCATGGTGTGATTCGGGACACCGCATTTGGCTACGAG 413
DB 376 GCGTCACTGCTTGCACCGCATGGTGTGATTCGGGACACCGCATTTGGCTACGAG 435
QY 414 ACCGTCAACACCGCCGGAAGGCGCACTACCTGCTGCCCCGATGAGCGAAGTCGCGGT 473
DB 436 ACCGTCAACACCGCCGGAAGGCGCACTACCTGCTGCCCCGATGAGCGAAGTCGCGGT 495
QY 474 CGACTGCGCGCCAGGTTGGCGCTTACCACTGATTCGGAACCCAGGGGCGCGGTGTG 533
DB 496 CGACTGCGCGCCAGGTTGGCGCTTACCACTGATTCGGAACCCAGGGGCGCGGTGTG 555
QY 534 CTGATGGGCGGGGTGCGGCGCTGCAACCGGCGGACGCTGCTGATTCGGGCGCGGAC 593
DB 556 CTGATGGGCGGGGTGCGGCGCTGCAACCGGCGGACGCTGCTGATTCGGGCGCGGAC 615

```


QY 894 CGACCGACACCTACGACACCGACCTTCGCGTCGACGACACGCTGTTTACTGCGTG 953
 |||||
 DB 916 CGACCGACACCTACGACACCGACCTTCGCGTCGACGACACGCTGTTTACTGCGTG 975
 |||||
 QY 954 GCGAACATGCCCCGCTTCGCGACGACGCTTCGACCTACCGCTGACCAACGCGACGATG 1013
 |||||
 DB 976 GCGAACATGCCCCGCTTCGCGACGACGCTTCGACCTACCGCTGACCAACGCGACGATG 1035
 |||||
 QY 1014 CCGATATGCTGAGCTTCGCGACGACGCTTCGCGTCGCGCTGCGCTGCAATCCGCGA 1073
 |||||
 DB 1036 CCGATATGCTGAGCTTCGCGACGACGCTTCGCGTCGCGCTGCGCTGCAATCCGCGA 1095
 |||||
 QY 1074 CTAGCCAAAGGCTTTTCGACGACGAGAGGCGCTTACTGTCGAAACGCTGCGCCACCGAC 1133
 |||||
 DB 1096 CTAGCCAAAGGCTTTTCGACGACGAGAGGCGCTTACTGTCGAAACGCTGCGCCACCGAC 1155
 |||||
 QY 1134 CTGGGGGTGCGCTTCACGCGACCGCGCGAGCGTGTGCGCTGACCTTCGCGCGCTTAC 1193
 |||||
 DB 1156 CTGGGGGTGCGCTTCACGCGACCGCGCGAGCGTGTGCGCTGACCTTCGCGCGCTTAC 1215
 |||||
 QY 1194 GCGGACGACGCTGCGAGTAAAGGAGCGATGATG 1229
 |||||
 DB 1216 GCGGACGACGCTGCGAGTAAAGGAGCGATGATG 1251
 |||||

RESULT 4
 ID V49511 standard; DNA; 682 BP.
 AC V49511:
 DT 20-OCT-1998 (first entry)
 DE Mycobacterium marinum Mar3 DNA.
 KW Alantine dehydrogenase; ADH; diagnosis; tuberculosis; pathogen;
 KM swimmers disease; vaccine; epidemic; infection; identification; ss.
 OS Mycobacterium marinum.
 FH Key Location/Qualifiers
 FT CDS 1..682
 FT /*tag= a
 FT /codon_start= 3
 FT /product= "Mar3"
 FT /note= "Alantine dehydrogenase"

W09832862-AZ.
 PD 30-JUL-1998.
 PF 29-JAN-1998; E00484.
 PR 29-JAN-1997; EP-101339.
 PA (FLOH/) FLOHE L.
 PI Flohe L, Hutter B, Kolik A, Singh M;
 DR MPI; 98-427958/36.
 DR P-PSDB; W64481.
 PT Nucleic acid encoding alantine dehydrogenase of Mycobacterium marinum
 PT used for diagnosis of tuberculosis and other mycobacterial
 PT diseases, also for treatment and prevention, for drug screening and
 PT for bio-transformation.
 PS Claim 1: Page 34-35; 57pp; German.
 CC This sequence encodes an alantine dehydrogenase (ADH) protein, Mar3
 CC isolated from Mycobacterium marinum. This protein is used to diagnose
 CC tuberculosis and other mycobacterial infections (including 'swimmers
 CC disease', caused by M. marinum, a fish pathogen) in humans or animals.
 CC The protein can also be used for control of epidemics and for
 CC vaccination, to screen for agents with anti-mycobacterial activity, and
 CC in bio-transformations that are specific for L-alanine. Also mycobacteria
 CC can be identified by analysis of genomic ADH sequences. ADH is an antigen
 CC that is secreted early during infection.
 CC Sequence 682 BP; 105 A; 254 C; 225 G; 98 T;

Query Match 36.9%; Score 454; DB 1; Length 682;
 Best Local Similarity 80.1%; Pred. No. 4.6e-86;
 Matches 546; Conservative 0; Mismatches 135; Indels 1; Gaps 1;

QY 98 AATTCGCGGTGCGACACCGCGCGCTCGCGAATACCGCTGCGCATGAG 157
 |||||
 DB 1 AATTCGCGGTGCGATCACCGCGCGCTCGCGCTGACCAACGCGCGACGAG 60
 |||||

QY 158 TGCTATCCAGGAGGTGCGGAGAGGCGCTACGACGCGGATTCAGAGGCG 217
 |||||
 DB 61 TGCTATCCAGGAGGTGCGGAGAGGCGCTACGACGCGGATTCAGAGGCG 120
 |||||
 QY 218 CAGGCGCGAATGCTGCGGACCGCGGACGAGTGTGGGCGGACGCTGATTTATGCTCA 277
 |||||
 DB 121 CCGGTGCCAGCTGATCAGACCGCGGACGAGTGTGGGCGGACGCTGCTCA 180
 |||||
 QY 278 AGGTCAAAAGACCGATAGCGCGGGAATACGCGCGCTGCGACAGGGC- GATCTGTTCA 336
 |||||
 DB 181 AGGTCAAAAGACCGATAGCGCGGGAATACGCGCGCTGCGGCGGCGGACGCTGTTCA 240
 |||||
 QY 337 CGTTCTGATTTGGGCGGCGTACGCTGCTGACCGATGCTGTTGATTCGGGACCA 396
 |||||
 DB 241 CCGTACTGACCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCA 300
 |||||
 QY 397 CGTCAATGCTTACGAGACGCTGCGAGCGCGGACGCGGACGCTGCGCGCA 456
 |||||
 DB 301 CGTCAATGCTTACGAGACGCTGCGAGCGCGGACGCGGACGCTGCGCGCA 360
 |||||
 QY 457 TGAAGCAAGTCCGCGCTGCGACTGCGCGCGCGGAGTGTGGCTTACCACTGATGCAAGCC 516
 |||||
 DB 361 TGAAGCAAGTCCGCGCTGCGACTGCGCGCGCGGAGTGTGGCTTACCACTGATGCAAGCC 420
 |||||
 QY 517 AAGGGGCGCGGCTGCTGATGCGGCGGCGCGGCGCTGCAACGCGCGAGCTGTAG 576
 |||||
 DB 421 ACGGGGCTGCGGCGCTGCTGATGCGGCGGCGCTGCGCGCTGCAACGCGAGCTGTAG 480
 |||||
 QY 577 TGATGCGCGCGCGGACCGCGGCTGCAACGCGCGGATGCGCGAGCTGTAGCGGA 636
 |||||
 DB 481 TGATGCGCGCGCGGACCGCGGATGCAACGCGCGGCTGCGCGAGCTGTAGCGGA 540
 |||||
 QY 637 CCGTTACGCTTACGACATCAACATGCGGACGCTGCGAGCGGAGTGTAG 696
 |||||
 DB 541 TGCTACCGCTGCTGATGCTCAATCAATCAACGCTGCGCGAGTGTAGCGG 600
 |||||
 QY 697 GCGGATCCACACTGCTACTGCTGCGGCTGCGAGGCTGCGGAGCTGCAACGCTGCGG 756
 |||||
 DB 601 GTCGCTCCGAGACCGCTGCTGCTGCGGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGG 660
 |||||
 QY 757 ACTGTGATTTGGGCGGCTGCT 778
 |||||
 DB 661 ACATGTGATGCGGCGGCTGCT 682
 |||||

RESULT 5

ID N91423 standard; DNA; 1125 BP.
 AC N91423:
 DT 01-FEB-1991 (first entry)
 DE Sequence of heat-resistant alantine dehydrogenase (AH) gene with mol. wt.
 DE 2 MD or less
 KW Enzyme; ds.
 OS Bacillus stearothermophilus IFO 12550.
 PN J01043194-A.
 PD 15-FEB-1989.
 PF 10-AUG-1987; 200524.
 PR 10-AUG-1987; JP-200524.
 PA (NIRA) Unilika KK.
 DR MPI: 89-096095/13.
 PT Recombinant plasmid for transforming Escherichia coli -
 PT obtl. by connecting heat resistant alantine dehydrogenase gene to
 PT vector plasmid
 PS Disclosure: Fig 2, p693; 9pp; Japanese.
 CC A recombinant plasmid contg. heat-resistant alantine dehydrogenase (AH)
 CC gene with mol. wt. 2 or less MD is claimed. Cells transformed with the
 CC vector produce high levels of heat-resistant AH. Transformed E. coli
 CC is useful as a clinical inspection reagent.
 CC Sequence 1125 BP; 246 A; 295 C; 367 G; 217 T;

Query Match 21.6%; Score 265.2; DB 1; Length 1125;
 Best Local Similarity 56.1%; Pred. No. 5.5e-47;


```

QY 360 CGTGTTCACCGATGCTGTGTGATTCGGCACCACGATTCGCTACGAGACGCTC 419
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 21414 CCAGATTTAGCAGATGCTATGTT-----AACGAGAAAAAACATGAACTGTT 21367
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 420 CAGACCGCGGAGGCGCATACCCCTGCTTCCCGCATAGCGAGAGTCCCGGTGACTC 479
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 21366 CGTGACATCAAGGAGACATACCGCTCCCTGCTCCATAGTGAAGTGCAGCTGTATG 21307
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 480 GCGGCCGAGGTTGGGCTTACCGATGATGCGAACCCAGGGGCGCGGTGGCTGATG 539
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 21306 GCTGTTCATTAATGAGAGCTACTTCTTACTTAAGCAAGCTGGTGGCTTGTCTTACTT 21247
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 540 GCGGGGCGCGCGCGCTGCAACCGCGGAGCTGTGTGATGCGCGCGCGCGCGCG 599
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 21246 GGGGGGTACAGAGTGTTCCTCAAAAAGGAAAGTAACTATCATCGGTGTGTGCTGCT 21187
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 600 TACACGCGACCGCGCATGCGCAAGCGCATGGCGCGCATCGTTTACATCATCAAC 659
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 21186 ACACATGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 21127
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 660 ATGCAACAACTTCGCGCACTGCGACCGGAGTTCGCGCGGATCCACACTGCTACTCA 719
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 21126 TCGAAGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 21067
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 720 TCGGCTTACGAGCTGAGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 779
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 21066 AATTCATTAACATTTGACCAAGTGAAGATGCTGATGCTGATTTGAGACCAATTC 21007
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 780 GTGCGAGGCGCGCAAGGCGCAACCAATGATGCTGATGCTGCTGCTGCTGCTGCT 21007
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 21006 ATCCCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 20947
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 840 GGTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 896
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 20946 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 20888
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 897 CGGACCACTTACGAGCACCAGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 956
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 20887 GTGACCAACCGAGATGACCCGCTGATGAAACACGAGTGTCTCCACTATCCCTGCT 20828
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 957 AACATCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1016
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 20827 AATATCCCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 20768
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1017 TATGCTGCTGAGCTTGGCAGCATGGCT 1044
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 20767 TATATCGAAGCTTGGCTGGCAAAAGAT 20740
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 7

V74442/c
ID V74442 standard; DNA; 9280 BP.
AC V74442;

DE 16-MAR-1999 (first entry)
DR Staphylococcus aureus contig SPQ ID #131.

KM Computer readable medium; vaccine; S aureus infection; immunodetection;
KM cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KM skin infection; surgical wound infection; scalded skin syndrome; ds.
KM Staphylococcus aureus.

OS Key Location/Qualifiers
FH 841..900
FT misc-feature

FT /tag- a
FT /note- "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"

FT misc-feature 2641..2700
FT /tag- b
FT /note- "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"

```

FT misc-feature 4441..4500
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
FT /tag- c
FT /note- "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT misc-feature 6241..6300
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
FT /tag- d
FT /note- "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT misc-feature 8041..8100
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
FT /tag- e
FT /note- "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT EP-786519-A2.
FT 30-JUL-1997.
FT 07-JAN-1997: 100117.
FT 05-JAN-1996: US-009861.
FT (HMA-) HUMAN GENOME SCI INC.
FT Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA,
FT Rosen CA;
FT WPI: 97-374922/35.
FT Polynucleotide(s) and proteins derived from Staphylococcus aureus
FT stored on computer readable medium and used in the production of
FT anti-S aureus vaccines
FT Claim 1; Page 710-715; 3271pp: English.
FT This sequence represents one of 5191 Staphylococcus aureus DNA sequences
FT of the invention. The DNA sequences are recorded on a computer readable
FT medium, preferably selected from a floppy or hard disk, random access
FT memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
FT the S aureus DNA sequences allows putative functions to be assigned so
FT that protein encoding or regulatory regions of commercial, therapeutic or
FT industrial importance can be obtained. Specifically, sequences which are
FT likely to encode antigens have been identified and these polypeptides can
FT be used in a vaccine composition against S aureus infection. The
FT CC polypeptides can also be used in a kit for the immunodetection of
FT S aureus in a sample. S aureus is implicated in numerous human diseases,
FT including cellulitis, eyelid infections, food poisoning, osteomyelitis,
FT skin and surgical wound infections, scalded skin syndrome, toxic shock
FT syndrome, etc. Organisms transformed with the DNA sequences can be used
FT for recombinant production of the polypeptides. The new DNA sequences
FT (and their fragments) are useful as primers or probes for isolating
FT homologues of any of the S aureus DNA sequences contained on the
FT computer readable medium.
FT Sequence 9280 BP: 2958 A; 1649 C; 1309 G; 3060 T;

```

Query Match 13.5W; Score 156.4; DB 1; Length 9280;
Best Local Similarity 50.5%; Pred. No 2.3e-26;
Matches 483; Conservative 0; Mismatches 466; Indels 7; Gaps 3;

```

QY 151 CATGAGTGTCTATCCAGCGAGTGGCGGAGAGGCTGCGCTATCACCGAGGATTC 210
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9280 CATACTGTTTATGAGAAACATGCGGTTACAGATCATCTTTGAAAGATGATATAC 9221
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 211 AAGCGCGAGCGCGCACTGCTGCGCAGCGCGGACCAAGTGTGGCGCGAGCTGATTA 270
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9220 AAGAGAGAGGTGCTGAGATTTGCTGAACAAACAAAGTTGG---GATGTGATATG 9164
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 271 TTGCTCAAGTCAAGAAGACGATGCGCGGAAATAGGCGCGCGGACAGCGGAGC- GATC 329
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9163 GTATTAAAGTAAAGAAACCACTGTAATCTGAATATCTTTAAAGAAAGGCTGTGA 9104
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 330 TTGTTCAGTCTTTCATTTGGCGCGCTCAAGTGTGCAACCGATGCTGTGTGATTC 389
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9103 TTATTCATTTATCTTCAATAGCAATAGAAAGAAATTAACAGCTTGTATAGATAG 9044
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 390 GGCACACGTCATGCTTACGAGACCGTCCAGACCGCGGAGGCGGACATCCCTGCTT 449
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

DB 9043 AAGTATTAGTATTGATATGAGACTGTGACATTAACACCGCATCTTACCATTTGTTA 8984
 QY 450 GCGCCGATGAGCAAGCGCGGATGACCTGCGCCGAGGTGGCGTTTCCACCTGATG 509
 DB 8983 TCACCATGATGATGAGGAGGAGGAGGATTCAGCTCAAGTTCAGAGAGTTCTACAA 8924
 QY 510 CGAACCAGAGGAGGCGCGGTGTGATGAGGGGGTGGCGCGTGAACCGCGGAC 569
 DB 8923 AACTTAATGATGATGAGGAGGAGGATTCAGTGTGATGCTCCAGAGACTTAAGGGTAA 8864
 QY 570 GTGCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 629
 DB 8863 GTACTATTTATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8804
 QY 630 GGGGCGAGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 689
 DB 8803 GGGGCGAGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8744
 QY 690 TTCTGCGCGGATTCACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 749
 DB 8743 TTCTGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8684
 QY 750 CGTCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 809
 DB 8683 CAAAGTATTTAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8624
 QY 810 TCAGATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 869
 DB 8623 ACAGAGACATGATTAACAAATGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8664
 QY 870 CAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926
 DB 8563 CAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8504
 QY 927 GTGACGACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 966
 DB 8503 AAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8444
 QY 987 ACCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1046
 DB 8443 ACCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8384
 QY 1047 CCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1102
 DB 8383 AGAGAGCATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8328

RESULT 8
 T67971
 ID T67971 standard: DNA: 1074 BP.
 AC T67971,
 DT 15-JUL-1997 (first entry)
 DE H. pylori membrane protein ORF 05cp20518orf61.
 KW Vaccine; prevention; treatment; infection; identification;
 KM binding compound; bacterium; life cycle; activator; bacteria;
 QM inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis;
 KM membrane; amino acid; metabolism; de.
 QM Helicobacter pylori.
 FH location/Qualifiers
 key 1.1074
 FT /tag a
 FT /note "no stop codon given"
 PN W09640893-A1.
 PD 19-DEC-1996.
 PE 06-JUN-1996; U09122.
 PR 07-JUN-1995; US-487032.
 PR 01-APR-1995; US-630405.
 PA (ASTR) ASTRA AB.
 PI Berglinch OT, Smith D, Mellgaard BL;
 DR MPI: 97-052306/05.
 DR P-PSDB: M20718.
 PT Helicobacter pylori nucleic acid sequences and related
 PT polypeptides) - useful for vaccines to treat or prevent H. pylori

PT Infection, and to detect Helicobacter
 PS Claim 27: Page 819: 1481pp: English.
 CC The present sequence encodes a Helicobacter pylori membrane
 CC protein likely to contain four membrane spanning regions.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds,
 CC useful as potential H. pylori life cycle activators or inhibitors.
 CC The genomic sequence of H. pylori (ATCC 55679) was determined from
 CC overlapping contigs generated by mechanically shearing the
 CC bacterial DNA. The sequences were analysed for ORF of at least 180
 CC nucleotides, and the predicted coding regions defined by computer
 CC evaluation. To identify likely H. pylori antigens for vaccine
 CC development, the amino acid sequences predicted from various ORF
 CC were analysed for significant homology to other known or exported
 CC membrane proteins. Having identified and determined the sequences
 CC of interest, particular regions can be isolated from H. pylori by
 CC PCR amplification for recombinant polypeptide production, e.g. in
 CC E. coli hosts.
 SQ Sequence 1074 BP; 302 A; 189 C; 297 G; 286 T;

Query Match 7.3%; Score 89.2; DB 1; Length 1074;
 Best Local Similarity 46.5%; Pred. No. 2e-10;
 Matches 373; Conservative 0; Mismatches 413; Indels 16; Gaps 2;

QY 265 GATTATTTGCTCAGGCTCAAGACCGATGAGGCGGAGTATGAGGCGGCTGCGA-CAGGG 323
 DB 133 GATTGTGCTGCTCAATGCAAGAGGCTTTAGAGCATGATTAACCTTTGCTCAAGAAAA 192
 QY 324 GCGATCTGCTGCTCAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 383
 DB 193 GCGACTGCTTTAGTATTTAGTATTTAGGCTATCAAAAAGCTTGTAATGTTTAT 252
 QY 384 GATTCGCGACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 443
 DB 253 AATTAATAATATCTCTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 312
 QY 444 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 495
 DB 313 ATTTTAGCGCTTATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 372
 QY 495 -----GCTTACACCTGATGCGAACCAAGGCGCGGCTGCTGCTGCTGCTGCTGCTG 548
 DB 373 TTACTGCTTTAGAGATTTAAAGGTTTATGGTAAGAGGGGCTGCTAGGGGTTTG 432
 QY 549 CCGGCGCTGAAACCGGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 608
 DB 433 TCGGCTGCGCAAGGCTTAAATCGCTAATTTGAGGCGGCTGCTGCTGCTGCTGCTGCT 492
 QY 609 GCGCGATGCGCAACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 668
 DB 493 GCGAAGCTTTAAGCAATGAGGCTTAAAGTACAGCTTTTGAATTTAGCTACCTTAA 552
 QY 669 CTTCGCAACTGACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 728
 DB 553 TTACAAACCCACCTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 612
 QY 729 GAGCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 788
 DB 613 AATATCATTCAAGCTTAAACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 672
 QY 789 GCGAAGGCAACCAATTTAGTCTGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 848
 DB 673 AGCAAAACCCCTTAAGTATTTAAGAGGCAATTTAAATATGATGACAGCAAGGGGTA 732
 QY 849 CTGCTGATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 908
 DB 733 GTCAATTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 792
 QY 909 GACCAACCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 968
 DB 793 TCTAACCGGCTTATGCTGAGAAAGTTTGTGCTATTTAGGCTGCTGCTGCTGCTGCTGCT 852


```

QY 969 TCGGTCCGAGACGTCGACCTACGCGCTGACCAAGCGCAGCATGCCGTATGCTCGAG 1028
   || || || || || || || || || || || || || || || || || || || || ||
Db 853 ATTGTGCTTAAACGAGCGCTTACGCGCTTATAGCCATGCGAGTGTCCGTATTGTGTAT 912
QY 1029 CTTCGCGACCATGCTGCGGG 1050
   || || || || || || || || || || || || || || || || || || || || ||
Db 913 TATTAGACATGCGCTTGAAG 934

RESULT 9
V43039/c
ID V43039 standard; DNA; 544 BP.
AC V43039;
DE 09-NOV-1998 (first entry)
KW Streptococcus pneumoniae polypeptide coding region;
KM Polypeptide; ORF; open reading frame; infection; bacterial;
OS Streptococcus pneumoniae; bacteremia; diagnosis; prophylaxis; ds.
FH Key Location/Qualifiers
FT CDS complement (59..334)
FT /tag= a
FT /note= "polypeptide"

W09823631-AL.
PD 04-JUN-1998.
PF 24-NOV-1997; U21976.
PR 27-NOV-1996; US-031879.
PA (SMK ) SMITHKLINE BEECHAM CORP.
PI (SMK ) SMITHKLINE BEECHAM PLC.
PI Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO,
PI Reid RH, Zarlos PN;
PI WPI: 98-322654/28.
PI P-PSDB: W62769.
DR Streptococcus pneumoniae polynucleotides - useful for developing
PT products for diagnosis, prevention and treatment of infections e.g.
PT pneumonia, bacteremia, meningitis or endocarditis
PT Claim 1, Page 165-166; 181p; English.
PS The sequence is that of a Streptococcus polypeptide coding region.
CC The polypeptide can potentially be used for the diagnosis and
CC prevention of bacterial infections, especially SP infection.
CC It may be used for the treatment of diseases such as otitis media,
CC conjunctivitis, pneumonia, bacteremia, meningitis, sinusitis, pleural
CC empyema, endocarditis or infection of the cerebrospinal fluid.
SQ Sequence 544 BP; 149 A; 121 C; 124 G; 150 T;

Query Match 5.6%; Score 69.2; DB 1: Length 544;
Best Local Similarity 53.4%; Pred No. 2.6e-06;
Matches 191; Conservative 0; Mismatches 163; Indels 4; Gaps 2;

```

```

RESULT 10
ID V4370
AC V4370 standard; DNA; 31096 BP.
DE 16-MAR-1999 (first entry)
KW Streptococcus aureus confg SEQ ID #59.
KM Computer readable medium; vaccine; S.aureus infection; immunodetection;
KM cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KM skin infection; surgical wound infection; scalded skin syndrome;
KM toxic shock syndrome; ds.
OS Streptococcus aureus.
FH Key Location/Qualifiers
FT misc_feature 1201..1260
FT /tag= a
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"

FT misc_feature 3001..3060
FT /tag= b
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"

FT misc_feature 4801..4860
FT /tag= c
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"

FT misc_feature 6501..6560
FT /tag= d
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"

FT misc_feature 8401..8460
FT /tag= e
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"

FT misc_feature 10201..10260
FT /tag= f
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"

FT misc_feature 12001..12060
FT /tag= g
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"

FT misc_feature 13801..13860
FT /tag= h
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"

FT misc_feature 15601..15660
FT /tag= i
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"

FT misc_feature 17401..17460
FT /tag= j
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"

```



```

FT misc-feature 19201..19260
FT /tag- k
FT /note- "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT
FT misc-feature 21001..21060
FT /tag- l
FT /note- "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT
FT misc-feature 22801..22860
FT /tag- m
FT /note- "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT
FT misc-feature 24601..24660
FT /tag- n
FT /note- "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT
FT misc-feature 26401..26460
FT /tag- o
FT /note- "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT
FT misc-feature 28201..28260
FT /tag- p
FT /note- "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT
FT misc-feature 30001..30060
FT /tag- q
FT /note- "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT
FT
FT EP-786519-A2.
FT 30-JUL-1997.
FT 07-JAN-1997: 100117.
FT 05-JAN-1996: US-009861.
FT (HUMA-) HDMA GENOME SCI INC.
FT Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA,
FT Rosen CA;
FT WPI: 97-374922/35.
FT Polynucleotide(s) and proteins derived from Staphylococcus aureus
FT stored on computer readable medium and used in the production of
FT anti-S.aureus vaccines
FT Claim 1: Page 452-469: 3271pp: English.
FT This sequence represents one of 5191 Staphylococcus aureus DNA sequences
FT of the invention. The DNA sequences are recorded on a computer readable
FT medium, preferably selected from a floppy or hard disk, random access
FT memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
FT the S.aureus DNA sequences allows putative functions to be assigned so
FT that protein-encoding or regulatory regions of commercial, therapeutic or
FT industrial importance can be obtained. Specifically, sequences which are
FT likely to encode antigens have been identified and these polypeptides can
FT be used in a vaccine composition against S.aureus infection. The
FT polypeptides can also be used in a kit for the immunodetection of
FT S.aureus in a sample. S.aureus is implicated in numerous human diseases,
FT including cellulitis, eyelid infections, food poisoning, osteomyelitis,
FT skin and surgical wound infections, scalded skin syndrome, toxic shock
FT syndrome, etc. Organisms transformed with the DNA sequences can be used
FT for recombinant production of the polypeptides. The new DNA sequences
FT (and their fragments) are useful as primers or probes for isolating
FT homologues of any of the S.aureus DNA sequences contained on the
FT computer readable medium.
FT Sequence 31096 BP; 11857 A; 5243 C; 5477 G 7488 T;
SQ

```

```

Query Match 4.6%; Score 56.8; DB 1: Length 31096;
Best Local Similarity 48.2%; Pred. No. 0.0013;
Matches 160; Conservative 0; Mismatches 172; Indels 0; Gaps 10;
QY 713 CTACTCATGGCCCTACGACCTGAGGCTCCCTCAACCTGCCGACCTGATTGGGCG 772
DB 649 CAATATCAACAGAAATTTAGCAGAACCAATTAAGAAAGATGATTTATTCTAC 708
QY 773 CGCTCGGTGCGAGGCGCCAGCAACCCCAATTAGCTCGAATTAATCTAGTGGCGAT 832
DB 709 AATTATTTTACGTGCGCAACCCCAAAATTTGTTACTGCTGAGATGTTAAATCAAT 768
QY 833 GAAACGAGTGGCGTACTGCTGATATACCATGACAGGCGGCTTTTCGAGGCTC 892
DB 769 GAAAGAGGTAGTATGATATACCTATGACCAAGGTGGACATTTGAAACAT 828
QY 893 ACGACCGACCTACGACACCCAGCTTCCCGTGCAGCAGACGCTGTTTACTGCGT 952
DB 829 TACACCACTACAAATTTCTGATCCAGTATGACAGAGGTGTGATTCATTATGCTGT 888
QY 953 GCGGACATGCCCGCTCGTCCGAGACGTCGACCTACGCGCTGACCAAGCGACGAT 1012
DB 889 ACCAATCAACGAGGACGTCACAGAACTTCAACATGTGATTAACCAAGAAATAT 948
QY 1013 GCGGTATGCTCGACCTTCCGACCATGCT 1044
DB 949 TGATTTATTTAGAAATTTTGACAAAGCT 980

RESULT 11
Q21833/C
ID Q21833 standard; DNA; 390 BP.
AC Q21833:
DE 08-JUN-1992 (first entry)
KW Randomising oligonucleotide used in SPERT mRNA prepn.
KW Systematic polypeptide evolution by reverse translation; SPERT;
OS Synthetic.
PN WO9202536-A.
PD 20-FEB-1992.
PF 01-AUG-1991: U05463.
PR 02-AUG-1990: US-561968.
PA (COLS ) UNIV OF COLORADO.
PI Gold L, Turk C;
DR WPI: 92-080018/10.
PT New method of systematic polypeptide evolution by reverse
PT translation - by linking each polypeptide in sample mixt. to
PT individualised mRNA allowing further synthesis of selected
PT polypeptide(s)
PS Example: Page 55: 102pp: English.
CC The sequence is that of an example randomising oligonucleotide which
CC is used in the prepn. of mRNA encoding candidate polypeptides for the
CC method of systematic polypeptide evolution by reverse translation
CC (SPERT). The method provides a rapid way of isolating and identifying
CC polypeptide ligands which bind to target moles. The polypeptide ligands
CC can be used in a g. assay method. Disentangling presences, cell sorting,
CC as activators or inhibitors of target mol. function, as probes, as
CC sequestering agents, drug delivery vehicles, modifiers of hormone
CC action and as catalysts. See also Q21830-Q21832.
SQ Sequence 390 BP; 125 A; 126 C; 133 G; 6 T;

Query Match 4.0%; Score 49.2; DB 1: Length 390;
Best Local Similarity 45.9%; Pred. No. 0.036;
Matches 168; Conservative 0; Mismatches 198; Indels 0; Gaps 0;
QY 425 CGCGGAAGCGCACTACCCCTGCTGCCGATGAGCAAGTCGCGGTCGACGCGCG 484
DB 366 CGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGT 307
QY 485 CCAGTTGGCGCTTACCACTGATGGAACCAAGGCGCGCGGTGTCGATGCGGCG 544

```


Fri Jun 23 09:31:36 2000

us-09-362-485-7.rng

Page 12

Search completed: June 22, 2000, 15:10:02
Job time: 17928 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 15:11:10 ; Search time 341.15 Seconds
(without alignments)
468.274 Million cell updates/sec

Title: US-09-362-485-7

Perfect score: 1229
Sequence: 1 ATCTTCAGATTAATCGAAC.....GAGTAAGGAGGATGATG 1229

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 segs, 6492525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/2/1na/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/1na/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/1na/5C.COMB.seq:*
4: /cgn2_6/ptodata/2/1na/5D.COMB.seq:*
5: /cgn2_6/ptodata/2/1na/6.COMB.seq:*
6: /cgn2_6/ptodata/2/1na/PCTUS.COMB.seq:*
7: /cgn2_6/ptodata/2/1na/Dackfil1es1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45.4	3.7	893	US-08-403-852D-3	Sequence 3, Appl1
2	45.4	3.7	5392	US-08-403-852D-1	Sequence 1, Appl1
3	44.4	3.6	2582	US-08-816-105A-2	Sequence 2, Appl1
4	43.6	3.5	28958	US-08-258-261B-6	Sequence 6, Appl1
5	43.6	3.5	28958	US-08-456-837-6	Sequence 6, Appl1
6	43.6	3.5	28958	US-08-457-342-6	Sequence 6, Appl1
7	43.6	3.5	28958	US-08-457-646A-6	Sequence 6, Appl1
8	43.6	3.5	28958	US-08-458-076A-6	Sequence 6, Appl1
9	43.6	3.5	28958	US-08-764-233A-4	Sequence 4, Appl1
10	43.6	3.5	28958	US-08-457-335A-6	Sequence 6, Appl1
11	43.6	3.5	28958	US-08-729-214-6	Sequence 6, Appl1
12	43.6	3.5	49377	US-08-764-233A-1	Sequence 1, Appl1
13	42.4	3.4	1833	US-08-403-852D-6	Sequence 6, Appl1
14	42.4	3.4	2185	US-08-173-508-3	Sequence 3, Appl1
15	42.4	3.4	2185	US-08-265-310-3	Sequence 3, Appl1
16	41.6	3.4	1620	US-08-461-775-10	Sequence 10, Appl1
17	41.6	3.4	20235	US-07-642-734C-3	Sequence 3, Appl1
18	40.6	3.3	4257	US-08-690-473-1	Sequence 1, Appl1
19	40.6	3.3	12001	US-08-458-568A-11	Sequence 11, Appl1
20	39.6	3.2	2414	US-08-461-775-11	Sequence 11, Appl1
21	39.6	3.2	2668	US-08-461-775-11	Sequence 11, Appl1
22	39.4	3.2	459	US-08-387-942C-35	Sequence 35, Appl1
23	39.4	3.2	1998	US-08-387-942C-35	Sequence 35, Appl1
24	39.4	3.2	12588	US-08-387-942C-1	Sequence 1, Appl1
25	39.2	3.2	4420	US-08-470-179-148	Sequence 148, App
26	39.2	3.2	44377	US-08-804-227C-7	Sequence 7, Appl1
27	39.2	3.2	44377	US-08-804-198-1	Sequence 1, Appl1

28	39	3.2	43280	US-08-804-227C-1	Sequence 1, Appl1
29	38.8	3.2	8051	US-08-576-626A-2	Sequence 2, Appl1
30	38.8	3.2	11219	US-07-642-734C-1	Sequence 1, Appl1
31	38.6	3.1	2064	US-08-343-428-1	Sequence 1, Appl1
32	38.6	3.1	11219	US-07-642-734C-1	Sequence 1, Appl1
33	38.4	3.1	1215	US-08-947-726A-1	Sequence 1, Appl1
34	38.4	3.1	2048	US-08-776-251-1	Sequence 1, Appl1
35	38.2	3.1	2109	US-08-555-568B-20	Sequence 20, Appl1
36	38	3.1	474	US-08-403-852D-14	Sequence 14, Appl1
37	38	3.1	30001	US-08-125-468-1	Sequence 1, Appl1
38	38	3.1	30001	US-08-474-933-1	Sequence 1, Appl1
39	37.8	3.1	3252	US-08-809-740A-1	Sequence 1, Appl1
40	37.8	3.1	3252	US-08-809-740A-4	Sequence 4, Appl1
41	37.8	3.1	15664	US-08-402-282-3	Sequence 3, Appl1
42	37.8	3.1	15664	US-08-508-004-3	Sequence 3, Appl1
43	37.8	3.1	15664	US-08-402-068-3	Sequence 3, Appl1
44	37.8	3.1	15664	US-08-402-068-3	Sequence 3, Appl1
45	37.6	3.1	1524	US-08-402-068-3	Sequence 3, Appl1
				Patent No. 5512669	

ALIGNMENTS

RESULT 1
US-08-403-852D-3
Sequence 3, Application US/08403852D
Patent No. 5891695
GENERAL INFORMATION:
APPLICANT: Blanc, Veronique
APPLICANT: Blanche, Francis
APPLICANT: Crouzet, Joel
APPLICANT: Jacques, Nathalie
APPLICANT: Lacroix, Patricia
APPLICANT: Thibaut, Denis
APPLICANT: Zagorec, Monique
APPLICANT: Debussche, Laurent
TITLE OF INVENTION: De Grety-Lagard, Valerie
TITLE OF INVENTION: Polypeptides Involved in The
TITLE OF INVENTION: Biosynthesis of Streptogramins, Nucleotide Sequences
TITLE OF INVENTION: Coding For These Polypeptides And Their Use
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Finnegan, Henderson, Farbow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/403,852D
FILING DATE: 10-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR 93/00923
FILING DATE: 25-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/11441
FILING DATE: 25-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03806.0054-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4400
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 833 base pairs
TYPE: nucleic acid


```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
COMPUTER: Apple Macintosh 6500
OPERATING SYSTEM: System 7.5
SOFTWARE: WordPerfect 3.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/816,105A
FILING DATE: 14-MAR-1997
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: DE 19625269.5
FILING DATE: 25-JUN-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: DE 19611252.4
FILING DATE: 22-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kurt G. Briscoe
REGISTRATION NUMBER: 33,141
REFERENCE/DOCKET NUMBER: Bayer 9814-KGB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914) 332-1700
TELEFAX: (914) 332-1844
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2587 nucleotides
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-816-105A-2

```

```

Query Match          3.6%; Score 44.4; DB 4; Length 2582;
Best Local Similarity 46.9%; Pred. No. 0.031;
Matches 138; Conservative 0; Mismatches 156; Indels 0; Gaps 0;

```

```

QY 352 CCGCGTCACGCTGCTTGCACGATGCGTTGTTGATTCGGCCAGCAGTCAATTCCTACG 411
    |||||  ||  |||||  ||  |||||  ||  |||||  ||  |||||  ||  |||||  ||
DB 1576 CCGCGTCACGCTGCTTGCACGATGCGTTGTTGATTCGGCCAGCAGTCAATTCCTACG 1635
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 412 AGACCGTCAGACGCGCGGAGGCGCACTACCTGCTTGCCTGATGAGGAGTGGCGG 471
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1636 TGACCGTCGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1695
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 472 GTGACATCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 531
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1696 GCGAGGTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1755
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 532 TGTGTATGGCGGGGGTCCCGGGGCTGCAACCGGCGGAGGAGGAGGAGGAGGAGGAGGAG 591
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1756 TCGGGCGGGTGTGACCGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 1815
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 592 CCGCGCGCTACAGCGAGCGGCGGATCGGCAAGCGGATGCGGCGGAGGAGGAGGAGGAGGAG 645
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1816 CCGCGCGGAGCGGTCGCGATCGGCGGAGTCCCGCGAGCGGTCGAGCACTGAGGAGGAG 1869
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

RESULT

4

```

US-08-258-261B-6
Sequence 6, Application US/08258261B
Patent No. 5639949

```

GENERAL INFORMATION:

```

APPLICANT: Schnupp, Thomas
APPLICANT: Lilgon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Gaffney, Thomas Deane
APPLICANT: Lam, Stephen Ting
APPLICANT: Hammer, Phillip E.
APPLICANT: Uknes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
antipathogenic substances
NUMBER OF SEQUENCES: 22

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532

```

COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/258,261B
FILING DATE: 08-JUN-1994
CLASSIFICATION: 800
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ. ID NO. 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOHETICAL: NO
ANTI-SENSE: NO
US-08-258-261B-6

```

```

Query Match          3.5%; Score 43.6; DB 1; Length 28958;
Best Local Similarity 44.8%; Pred. No. 0.096;
Matches 290; Conservative 0; Mismatches 354; Indels 4; Gaps 3;

```

```

QY 148 GGCATGAGTGTGCTATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 207
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 25306 GGCATGAGTGTGCTATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 25365
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 208 TTGAAGGCGGAGGCGGCGGAGTGTGCGACCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 267
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 25366 TTGAAGGCGGAGGCGGCGGAGTGTGCGACCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 25425
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 268 TTATGCTCAAGGTCAA-AGAACCGATAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 326
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 25426 TTATGCTCAAGGTCAA-AGAACCGATAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 25485
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 327 ATCTGTACGCTTCTTGTGATTTGGCGGCGGCTGACGCTTGTGACCGATGCTTGTGAT 386
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 25486 ATCTGTACGCTTCTTGTGATTTGGCGGCGGCTGACGCTTGTGACCGATGCTTGTGAT 25545
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 387 TCGGCGACGATGATTTGCTTACGAGAGCGGTCAGACCGGCGGAGGAGGAGGAGGAGGAGGAG 446
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 25546 TCGGCGACGATGATTTGCTTACGAGAGCGGTCAGACCGGCGGAGGAGGAGGAGGAGGAGGAG 25605
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 447 CTGCGCGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 504
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 25606 CTGCGCGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 25665
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 505 TGATGGAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 564
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 25666 TGATGGAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 25725
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 565 CCGAGC-TGTGTGTGTCGGCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 623
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 25726 CCGAGC-TGTGTGTGTCGGCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 25785
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```



```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/457,342
APPLICATION NUMBER: US/08/457,342
FILING DATE: 01-JUN-1995
CLASSIFICATION: 424
Prior Application DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/256,261
FILING DATE: 08-Jun-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
/JS-08-457-342-6

```



```

RESULT          9
US-08-764-233A-4
Sequence 4, Application US/08764233A
Patent No. 5716849
GENERAL INFORMATION:
APPLICANT: Ligon, James M.
APPLICANT: Schupp, Thomas
APPLICANT: Beck, James J.
APPLICANT: Hill, Dwight S.
APPLICANT: Neff, Snezana
APPLICANT: Ryals, John A.
TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSER: Ciba-Geigy Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/764,233A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/729,214
FILING DATE: 09-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: 1506/CIP6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Sorangium cellulosum
IMMEDIATE SOURCE:
CLONE: p98/1
US-08-764-233A-4

Query Match      3.5%; Score 43.6; DB 2; Length 28958;
Best Local Similarity 44.8%; Pred. No. 0.096; Indels 4; Gaps 3;
Matches 290; Conservative 0; Mismatches 354;

QY   148  GCCCATGA  CTCCTATCCAGGACAGGTGCCGGAGAGGC  CTCGGCTATACCAGCCGGAT  207
       |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db   25306  GCCCGAGCAAGCACCTCCGGATGGCCCTATTCGCTGCGCACACGCCGCCGCAC  25365

QY   208  TTCAGGGGGGCGAGCGCAACTGTGTGGGCACCGCGAACAGTGTGGGCGCAACCTGAT  267
       |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db   25366  TTCAGGCAACCGCGCGCTCTGCTAACCACAACCGCGAGACTCTCTCTCCGCGTCGAC  25425

QY   268  TTATGCTCAAGGTCAA-AGAACCAGTAAGCGCGGAGATACGGCCCTCTCGACACGGGCG  326
       |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db   25426  TCGGTCGGCCGAGGAAAGCCCCCGCCGAGCAACGTCCTCGAAGGAGGGAAGGCAACGGC  25485

Y    327  ATCTGTTCACAGTTCTTGATTTGGCCCGCGTCACGTGCTTCCACCGATGCGTTTGAT  386

```

Db	25486	AAAGTCGCTCTTCTGCTTTTCTTCTTGAGGCAAGGCTTGCAAGTGGGAAGGATGGGCTCTCTGCTG	254545
Qy	387	TCGCGACACACGTCATATTGCTTACGAGAACCGCTCCAGACCGCCGGAAGGCGCACTACCCCTG	446
Db	25546	CTCGACTCTCTGCGCCGCTCTTCGCGCGCTGAGCTGGAAGCATGCGAGCGCGGCGCTCGCTCT	25605
Qy	447	CTTCCGCCCATGAGCGAATTCGCGGTCGACTCTCGCCGCCCAAGGTTGGCGCTTACAC--C	504
Db	25606	CAGCTCGAGTGGAGCGCTGCTGCGCTCTGCGCCCGCGAGGAGGCGGCCCTCTCTTCGAC	25665
Qy	505	TGATGCCAACCCCAAGGGGGCGCGGCTGCTGATGGGGGGGCTCCCGCGCTCGAACCGG	564
Db	25666	CGCGTCGACGCTGTACAGGCGCCGCTCTTTCGCGTCAAGTGTCTCCCTGGCGGCCCTCTG	25725
Qy	565	CCGACG--TGTGTGATCGGCGCGGACCGCGGCTCTCAACGCAAGCCGCACTGCGCAAC	623
Db	25726	CGCTCGCTCGCGCTAGAGCGCGCGCGCGCTGCTGCGCCACATCTAGGGCGAGATCGCGCC	25785
Qy	624	GGCATGGCGCCGACCGCTTACGGGTTGACATCAATCAATGCAAACTTGGGCAACTCGAC	683
Db	25786	GCCTCTGCTCCAGGCGCTCTCTTCCTTCGAGAGACGGCGCCGCACTGCGCCCTGCGCAC	25845
Qy	684	GCGGAGTTCTGGCGGCGGATCCACACTGCTACTGCTACGAGTCTGAGGTCGCGGCTCC	743
Db	25846	AAACGCGTCAACACCGTCTCGCGCAACGGGGGCAATGGCGCGCTGAGTCTGGCGCTCC	25905
Qy	744	GTCAAAGTCCGCACTGCTGATTTGGGCGCGCTCTGGTGGCGAGCGCC	791
Db	25906	GACCTCGACACTACTCGCTCGCTGGGGGCGACAGGCTCTTCATCGCC	25953

REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8614
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 28958 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 US-08-457-335A-6

Query Match 3.5%; Score 43.6; DB 2; Length 28958;
 Best Local Similarity 44.8%; Pred. No. 0.096;
 Matches 290; Conservative 0; Mismatches 354; Indels 4; Gaps 3;

148 GGCCTAGAGTGTCTATCCAGGAGGCTCCGAGAGGCTCCGCTATCCAGGAGGAT 207
 25306 GCCCAGACACCTCTACCCCTCGGAGATGTGCTTATTCCTGACACACCCGCGCCAC 25365
 208 TTCAAGGCGGACGAGCGGCACTGTGCGACCGCGGACCGAGGTGTGGCGGAGCTGAT 267
 25366 TTGAGAGACCGGCGCCCTCTCTGATAGCCACACACCGGAGAGCTCTCTCCGCGTCGAC 25425
 268 TTATTGCTCAAGTCAAA-AGAACCGATAGCGGCGGAATAGCGCGCTCGGACAGCGGCG 326
 25426 TCAGTCGCGGACGAGCAAGCCCGCGGAGACCTCTCTCGGAGGAGCGGA] :CAGGCG 25485
 327 ATCTTGTCAGTTCAGGATTTGGCGGCGGCTGACGTGCTTGACCGATGGTGTGGAT 386
 25486 AACCTGCTCTCTCTCTCTCTCTGCGCAAGGCTCCAGTGGGAAGGATGGCCCTCTGCTG 25545
 387 TCCGACACACGATCAATTTGCTTACGAGACCGTCCAGACCGCGGAGCGGACATACCCCTG 446
 25546 CTCGACTCTCCGCGCCCT 25605
 447 CTGCGCGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 504
 25606 CAGCTGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 25665
 505 TGATGGGAACCCAAAGGGGGGCGGCGGTGTGCTGATGGGCGGCGGCGGCGGCGGCGGCG 564
 25666 CCGCTGAGAGCTGACAGCGCGCGCGCTTTGCGCGTATGATGATGATGATGATGATGAT 25725
 565 CCGAGC-TGCTGTGATTCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 623
 25726 CCGCTGCGCTGAGAGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 25785
 624 GGCATGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 683
 25786 GCCTTGTGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 25845
 684 GCGGATTTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 743
 25846 AAGCGCTTACACACCT 25905
 744 GTCAAGCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 791
 25906 GACCTGACAGCTACT 25953

RESULT 11
 US-08-729-214-6
 Sequence 6, Application US/08729214
 Patent No. 5817502
 GENERAL INFORMATION:
 APPLICANT: Ligon, James M.
 APPLICANT: Hill, Dwight Steven
 APPLICANT: Ryals, John Andrew

APPLICANT: Hammer, Phillip E.
 APPLICANT: van Pee, Karl-Heinz
 APPLICANT: Kirner, Sabine
 TITLE OF INVENTION: Genes for the synthesis of
 TITLE OF INVENTION: antipathogenic substances
 NUMBER OF SEQUENCES: 27
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Cliba-Gelby Corporation
 STREET: 520 White Plains Road
 CITY: Tarrytown
 STATE: NY
 COUNTRY: USA
 ZIP: 10591
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/729, 214
 FILING DATE: TBA
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Meigs, J. Timothy
 REGISTRATION NUMBER: 38,241
 REFERENCE/DOCKET NUMBER: CGC 1506/CIP5
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8587
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 28958 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 US-08-729-214-6

Query Match 3.5%; Score 43.6; DB 2; Length 28958;
 Best Local Similarity 44.8%; Pred. No. 0.096;
 Matches 290; Conservative 0; Mismatches 354; Indels 4; Gaps 3;

148 GGCCTAGAGTGTCTATCCAGGAGGCTCCGAGAGGCTCCGCTATCCAGGAGGAT 207
 25306 GCCCAGACACCTCTACCCCTCGGAGATGTGCTTATTCCTGACACACCCGCGCCAC 25365
 208 TTCAAGGCGGACGAGCGGCACTGTGCGACCGCGGACCGGAGGTGTGGCGGAGCTGAT 267
 25366 TTGAGAGACCGGCGCCCTCTCTGATAGCCACACACCGGAGAGCTCTCTCCGCGTCGAC 25425
 268 TTATTGCTCAAGTCAAA-AGAACCGATAGCGGCGGAATAGCGCGCTCGGACAGCGGCG 326
 25426 TCAGTCGCGGACGAGCAAGCCCGCGGAGACCTCTCTCTCTCTCTCTCTCTCTCTCTCT 25485
 327 ATCTTGTCAGTTCAGGATTTGGCGGCGGCTGACGTGCTTGACCGATGGTGTGGAT 386
 25486 AACCTGCT 25545
 447 CTGCGCGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 504
 25546 CTCGACT 25605
 505 TGATGGGAACCCAAAGGGGGGCGGCGGTGTGCTGATGGGCGGCGGCGGCGGCGGCGGCGG 564
 25666 CCGCTGAGAGCTGACAGCGCGCGCGCTTTGCGCGTATGATGATGATGATGATGATGATGAT 25725

QY 565 CCGAGC-TGCTGTGATTCGGCCGCGGACCGCGCTACAAACCGCCGATCGCCAC 623
DB 25726 CGCTCGCTGAGCGAGCCGCGCGCTGTCGCGACGAGGCGGATCGCGCGCC 25785
QY 624 GCGATGGCGGCGGCGCTTACGCTTACAGATCAACATCGACAACTGGCGCATCGAC 683
DB 25786 GCGTGTGCGGAGCGCTCTCTCCGAGAGCGCGCCGCGCTGCGCGCGACG 25845
QY 684 GCGGAGTCTGCGCGCGGATTCAGCTACGCTACGCTACGAGTACGAGGTCGC 743
DB 25846 AAGCGCTCAGCAGCGCTCGCGCGGACGCGCGCATGCGCGCGCTCGAGCTCGCGCGCTCC 25905
QY 744 GTCAACGTCGCGAGCTGTGTATGGGCGCGCTGCTGCGCGAGCGCC 791
DB 25906 GACCTCAGACCTACCTCGCTCGCGGCGGACGAGCTCTCATCGCC 25953

RESULT 12
US-08-764-233A-1
Sequence 1, Application US/08764233A
Patent No. 5716849

GENERAL INFORMATION:
APPLICANT: Ligoe, James M.
APPLICANT: Schupp, Thomas
APPLICANT: Beck, James J.
APPLICANT: Hill, Dwight S.
APPLICANT: Neff, Snezana
APPLICANT: Ryals, John A.
TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/764,233A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/729,214
FILING DATE: 09-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: 1506/CIP6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4937 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Sorangium cellulosum
IMMEDIATE SOURCE:
CLONE: p98/1, pJL3, and pVKM15
FEATURE:
NAME/KEY: misc_feature
LOCATION: 383..760

OTHER INFORMATION: /product= "SorR"
OTHER INFORMATION: /note= "this gene encodes a protein that is highly homologous to type I PKs such as eryA from Saccharopolyspora erythraea."
OTHER INFORMATION: Saccharopolyspora erythraea."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 927..19874
OTHER INFORMATION: /product= "Sora"
OTHER INFORMATION: /note= "Gene product is highly homologous to type I PKs t are known to be involved in the synthesis of polyketide compounds."
OTHER INFORMATION: compounds."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 942..7115
OTHER INFORMATION: /product= "Module 1 of Sora"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7203..12884
OTHER INFORMATION: /product= "Module 2 of Sora"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 13455..19616
OTHER INFORMATION: /product= "Module 3 of Sora"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19871..46318
OTHER INFORMATION: /product= "Sorb"
OTHER INFORMATION: /note= "Gene product is highly homologous to type I PKs ge
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19870..24556
OTHER INFORMATION: /product= "Module 1 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 24638..30820
OTHER INFORMATION: /product= "Module 2 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 30881..354
OTHER INFORMATION: /product= "Module 3 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 35528..40114
OTHER INFORMATION: /product= "Module 4 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 40190..46318
OTHER INFORMATION: /product= "Module 5 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 46851..47891
OTHER INFORMATION: /product= "Sorm"
OTHER INFORMATION: /note= "The protein encoded by the sorm gene is highly homologous to the methyltransferase from Streptomyces hygroscopicus that is involved in the synthesis of the polyketide rapamycin."
OTHER INFORMATION: polyketide rapamycin."
US-08-764-233A-1

Query Match 3.5%; Score 43.6; DB 2: Length 49377;
Best Loca. Similarity 44.8%; Pred. No. 0.11; Mismatches 354; Indels 4; Gaps 3;
Matches 290; Conservative 0;

QY 148 GCGCATGAGTCTCTATCCAGCGAGTCCGAGAGGCTGCTATACCGAGCGGAT 207
DB 41609 GCCACGAGCAGCTACCTCCGCGGATGTGCTATTCGTCGACACCGCGCGCAC 41668
QY 208 TTCAGGCGGCGGCGGCGCACTGTGCGGACCGCGGACGAGTGTGGCGGAGCTAT 267
DB 41669 TTCGAGCAGCGCGCGCTCTGTAGCCCAACCGCGGAGCAGCTCTCTCGCGCTGAC 41728
QY 268 TTATTCCTAAGGTCAA-AGAACCAGTAGCGGCGGAATACGCGCGCTGCGACAGGCGG 326

```

Db 41729 TCGCTGCCAGACAAAGCCGCCGAGACCGCTCTCGAGGAGCGGAAGCCAGGC 41788
Qy 327 ACTTTGTCAGCTTCTTGATTTGGCCGGCTACGTTGACCGATGCTTGTGAT 386
Db 41789 AAGCTGCTTCTTCTTCTTCTGCGCAAGCTCGCATGGGAAGATGGCCCTCTGCTG 41848
Qy 387 TCCGGACCAAGCTCAATTCCTTACGAGACCGTCCAGACCGCGGAGCGGACTTACCCCTG 446
Db 41849 CTCGACTCTCTGCGCCCTCTCTCCGCTCGAGCTGAGACATGAGCGCGCGCTCGCTCT 41908
Qy 447 CTGGCCCGATGAGCGAGTTCGCGCGCTCGACTGCGCGCGGATGGGCTTACCAAC--C 504
Db 41909 CACGTGATGAGAGCTGCTTCCGCTCTGCGCGCGGAGAGAGCGCGCCCTCTCTCGAC 41968
Qy 505 TGTATCGCAACCAAGGGGCGCGGCTGCTGATGAGCGCGCGGCTCGGAGCGG 564
Db 41969 CCGCTGAGCTGATGAGCGCGCGCGCTCTTGGCGTATGCTCTCCCTGCGCGCGCTCTG 42028
Qy 565 CCGAGC-TGCTGATGATGCGCGCGCGCGCGGCTACGAGCGCGCGGCGGCGGCGG 623
Db 42029 CCGCTGCGCGGCTGAGAGCGCGCGCGCGCTGCTGCGCGCGAGTACGAGTCCGCGCC 42088
Qy 624 GCGATGCGCGCGAGCTGATGCTTACGATACATCATC ACAAATTCGCGAGCTGAGC 683
Db 42089 GCGCTGCTGCGAGCGCGCTCTCTCCCTCGAGAGCGCGCGCGCGCTGCGCGAGC 42148
Qy 684 GCGGAGTCTGCGCGCGGATTCACACTGCTACTATGCGCGCGGCTGAGCTGAGCTGCG 743
Db 42149 AAGAGCTTACCAAGCTGCGCGCGCGCGAGCGCGCGCATGCGCGCGCTGCGCGCTCT 42208
Qy 744 GTCAACGTCGCGAGCTGCTGATGAGGCGCGCTGCTGCTGCGCGCGCGG 791
Db 42209 GACCTTCAGACCTGCTGCTGCTGCGCGCGCGAGCGCGCGCTGCTCATGGCC 42256

RESULT 13
US-08-403-852D-6
Sequence 6, Application US/08403852D
Patent No. 5891695
GENERAL INFORMATION:
APPLICANT: Blanc, Veronique
APPLICANT: Blanche, Francis
APPLICANT: Crouzet, Joel
APPLICANT: Jacques, Nathalie
APPLICANT: Lacroix, Patricia
APPLICANT: Thibaut, Denis
APPLICANT: Zagorec, Monique
APPLICANT: Debussche, Laurent
APPLICANT: De Crecy-Lagard, Valerie
TITLE OF INVENTION: Polypeptides Involved In The
TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences
TITLE OF INVENTION: Coding For These Polypeptides And Their Use
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Parabow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/403,852D
FILING DATE: 10-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR 93/00923
FILING DATE: 25-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/11441

```

```

FILING DATE: 25-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03806.0054-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4400
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1833 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: S.Pristinaespiralis
FEATURE:
NAME/KEY: CDS
LOCATION: 103..1689
US-08-403-852D-6

Query Match 3.48; Score 42.4; DB 3; Length 1833;
Best Local Similarity 50.5%; Pred. No. 0.087;
Matches 103; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

Qy 532 TGCATGAGGGGGGGTGGCCGCGCTCGACAGCGCGGAGCTGTGATGCGCGCGCA 591
Db 413 TGATCTGGAGCTGCTTCTTGGCGCGGCTCGCTGCGCGGCTTGGCCAGAGACT 472
Qy 592 CCGCGCGCTACACAGACGCGCGCATGCGCAAGCGCATGCGCGCGGCGGCGGCTTAG 651
Db 473 CCGCCCACTCATGCGCGCGCGCGCGCATGCGCGGCGGCGGCGGCGGCTGCGGAGCA 532
Qy 652 ACATCAATGACAAACTTGGCACTGACGCGCGGAGTTGCGCGCGGATCCACACTC 711
Db 533 CCACCTTGCCTGATCATGAGGCTCTTGCAGCAGAGCAAGCGCGCGCGGCGGATCGGC 592
Qy 712 GCTACTGCGCGCTACGA,CTCG 735
Db 593 TGTGGGAGCGCGCGGCGGCTGCG 516

RESULT 14
US-08-173-508-3
Sequence 3, Application US/08173508
Patent No. 5616485
GENERAL INFORMATION:
APPLICANT: Bartfield, Daniel
APPLICANT: Butler, Michael J.
APPLICANT: Hadary, Dany
APPLICANT: Jenish, David
APPLICANT: Krieger, Timothy
TITLE OF INVENTION: STREPTOMYCES PROTEASES AND IMPROVED
TITLE OF INVENTION: STREPTOMYCES STRAINS FOR EXPRESSION OF PEPTIDES AND
TITLE OF INVENTION: POLYPEPTIDES
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/173,508

```

```

FILING DATE: 23-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 18740/125/CACO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2185 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 531..2069
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 531..902
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 903..2069
FEATURE:
NAME/KEY: misc_feature
LOCATION: 531..533
OTHER INFORMATION: /note="Met at position -124"
OTHER INFORMATION: represents fmet"
US-08-173-508-3

```

```

Query Match          3.4%; Score 42.4; DB 1; Length 2185;
Best Local Similarity 47.4%; Pred. No. 0.091;
Matches 127; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

```

```

QY 524 CCGCGGTGCTGATGAGGGGGGTCGCCGGGTGACACCGCCGACGCGGTGATCGG 583
DB 1250 CGGCGTCCGATGACCGCGGTGACACCGCGGGTGAAGTGGCCGATCAAGTCTCAA 1309
QY 584 CGCGGACCGCGCGGTGACACCGCGGTGACACCGCGGGTGAAGTGGCCGATCAAG 643
DB 1310 CCGCGTCCGATGACCGCGGTGACACCGCGGGTGAAGTGGCCGATCAAGTCTCAA 1369
QY 644 GGTTCAGACATCAACATCAACAACTTCGGCACTCGACCGCGGTCTGCGCGCGAT 703
DB 1370 CGGCGTCCGATGACCGCGGTGACACCGCGGGTGAAGTGGCCGATCAAGTCTCAA 1429
QY 704 CCACACTCGTACTCATGCGCTTACGAGCTGAGAGGTGCGCTCAACGTCGCGACTGT 763
DB 1430 CGACCGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1489
QY 764 GATTGGGCGCTCTGCTGTCGACGCGCC 791
DB 1490 GAAGGCGCGGTGACGACGTCGCGCGCGCC 1517

```

```

RESULT 15
US-08-265-310-3
Sequence 3, Application US/08265310
PATENT INFORMATION:
PATENT NO. 5856166
GENERAL INFORMATION:
APPLICANT: Bartfeld, Daniel
APPLICANT: Butler, Michael J.
APPLICANT: Hadary, Dany
APPLICANT: Jenish, David
APPLICANT: Krieger, Timothy
APPLICANT: Malek, Lawrence T.
APPLICANT: Soostmeyer, Gisela
APPLICANT: Walczyk, Ewa
APPLICANT: Krygsmann, Phyllis
APPLICANT: Garven, Sheila

```

```

TITLE OF INVENTION: STREPTOMYCES PROTEASES AND IMPROVED
TITLE OF INVENTION: STREPTOMYCES STRAINS FOR EXPRESSION OF PEPTIDES AND
TITLE OF INVENTION: POLYPEPTIDES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/265,310
FILING DATE: 24-JUN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,508
FILING DATE: 23-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 18740/133/CACO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2185 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 531..2069
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 531..902
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 903..2069
FEATURE:
NAME/KEY: misc_feature
LOCATION: 531..533
OTHER INFORMATION: /note="Met at position -124"
OTHER INFORMATION: represents fmet"
US-08-265-310-3

```

```

Query Match          3.4%; Score 42.4; DB 3; Length 2185;
Best Local Similarity 47.4%; Pred. No. 0.091;
Matches 127; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

```

```

QY 524 CCGCGGTGCTGATGAGGGGGTCCCGGCTGCAACCGCGGACGTCGTGATCGG 583
DB 1250 CGGCGTCCGATGACCGCGGTGACACCGCGGGTGAAGTGGCCGATCAAGTCTCAA 1309
QY 584 CGCGGACCGCGCGGTGACACCGCGGTGACACCGCGGGTGAAGTGGCCGATCAAG 643
DB 1310 CCGCGTCCGATGACCGCGGTGACACCGCGGGTGAAGTGGCCGATCAAGTCTCAA 1369
QY 644 GGTTCAGACATCAACATCAACAACTTCGGCACTCGACCGCGGTCTGCGCGCGAT 703
DB 1370 CGGCGTCCGATGACCGCGGTGACACCGCGGGTGAAGTGGCCGATCAAGTCTCAA 1429
QY 704 CCACACTCGTACTCATGCGCTTACGAGCTGAGAGGTGCGCTCAACGTCGCGACTGT 763
DB 1430 CGACCGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1489

```

Fri Jun 23 09:31:37 2000

us-09-362-485-7.rni

Page 12

Oy 764 GATTGGGCGCTGTGCCAGGCGCC 791
||| ||| ||| ||| |||
Db 1490 GAAGGCGCGGTCAACGTGCGCGGCC 1517

Search completed: June 22, 2000, 15:13:36
Job time: 18294 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 12:07:43 ; Search time 5541.94 Seconds
(Without alignments)
898.857 Million cell updates/sec

Title: US-09-362-485-7
Perfect score: 1229
Sequence: 1 ATCTTGACATTAATCAAC.....GAGTAAGGACGATGATG 1229

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 202611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST: *
1: em_est1: *
2: em_est2: *
3: em_est3: *
4: em_est4: *
5: em_est5: *
6: em_est6: *
7: em_est7: *
8: em_est8: *
9: em_est9: *
10: em_est10: *
11: em_est11: *
12: em_est12: *
13: em_est13: *
14: em_est14: *
15: em_est15: *
16: em_est16: *
17: em_est17: *
18: em_est18: *
19: em_est19: *
20: gb_est1: *
21: gb_est2: *
22: gb_est3: *
23: gb_est4: *
24: gb_est5: *
25: gb_est6: *
26: gb_est7: *
27: gb_est8: *
28: gb_est9: *
29: gb_est10: *
30: gb_est11: *
31: gb_est12: *
32: gb_est13: *
33: gb_est14: *
34: gb_est15: *
35: gb_est16: *
36: gb_est17: *
37: gb_est18: *
38: gb_est19: *
39: gb_est20: *
40: gb_est21: *
41: gb_est22: *
42: gb_est23: *
43: gb_est24: *
44: gb_est25: *

45: gb_est26: *
46: gb_est27: *
47: gb_est28: *
48: gb_est29: *
49: gb_est30: *
50: gb_est31: *
51: gb_est32: *
52: em_est20: *
53: em_est21: *
54: em_est22: *
55: em_est23: *
56: em_est24: *
57: em_est25: *
58: em_est26: *
59: gb_est33: *
60: gb_est34: *
61: gb_est35: *
62: gb_est36: *
63: gb_est37: *
64: gb_est38: *
65: em_est27: *
66: em_est28: *
67: em_est29: *
68: em_est30: *
69: gb_est39: *
70: gb_est40: *
71: gb_est41: *
72: gb_est42: *
73: gb_est43: *
74: gb_est44: *
75: em_est31: *
76: em_est32: *
77: em_est33: *
78: em_est34: *
79: gb_est45: *
80: gb_est35: *
81: gb_est36: *
82: gb_est37: *
83: gb_est38: *
84: gb_est39: *
85: gb_est40: *
86: em_est34: *
87: em_est35: *
88: em_est36: *
89: em_est37: *
90: gb_est41: *
91: gb_est42: *
92: gb_est43: *
93: gb_est44: *
94: gb_est45: *
95: em_est40: *
96: em_est41: *
97: em_est42: *
98: em_est43: *
99: em_est44: *
100: em_est45: *
101: em_est46: *
102: gb_est47: *
103: gb_est48: *
104: em_est49: *
105: gb_est50: *
106: gb_est51: *
107: gb_est52: *
108: gb_est53: *
109: gb_est54: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES

OY	448	TTCGCCCCATGAGCGAAGTGCCTGGCTACTGACCTCCGCCAGGTGGCATTACCACTTA	507
Db	248	AAGMGSMSNGVGGCACAAAGAAGGGCGGAACCCTCCGCACASCCSSSCSSCSGCCSSSSS	189
OY	508	TTCGAACCCCAAAGGGGGCCCGCGGTGTCTGTATGGGGCGGGGTCCCGCGTTGAACGGCCG	567
Db	188	CCTGACCCCCSSAAGGGCCCCSGCGCCACCCCCTCCCGSAGCCVAGACCCCTCAC -CMWVGGSF	130
OY	568	ACGTCTGTGTGATTCGGCGCGCGGACCGCGCGGTATCAACAAGCAGCCCGCATCGCCAACGGCA	627
Db	129	ACGCAAAAGGGGAGAGAGGGCCGGNAGCCCCCGCNCACACCCCCCGCCGCACAGCCCGGGCA	70
OY	628	TGGGCGCGAC	638
Db	69	ACGCGCCAAC	59
RESULT 13			
LOCUS	CNS00720	932 bp	DNA
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC # BAC11409 of Rpci-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		
ACCESSION	AL066742		
VERSION	AL066742.1 GI:4945205		
KEYWORDS	GSS.		
SOURCE	fruit fly.		
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 932)		
REFERENCE	Genoscope.		
AUTHORS	Direct Submission		
TITLE	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr		
JOURNAL	- Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACS. For further information please see http://www.fruitfly.org/The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammotter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named Rpci-98 and was constructed by partial EcoRI digestion of Drosophila DNA prov^4-d by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters fo. hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.		
COMMENT	location/Qualifiers		
FEATURES	1..932		
Source	/organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone_lib="Rpci-98" /clone="BAC11409" /note="end : 17"		
BASE COUNT	155 a	202 c	241 g 91 t 243 others
ORIGIN			
Query Match	3.6%	Score 44.8;	DB 82; Length 932;
Best Local Similarity	34.1%;	Pred No. 15;	
Matches 73;	Conservative 47;	Mismatches 94;	Indels 0; Gaps 0;
OY	432	GAGCGACTACCCCTCTTGCCCGATGCGAAGTGCAGCGCGTCTGACTCGCCGCCAGGTT	491
Db	679	GCSGCGCGCGCCCCSSGSCCGCGCGSCGSGCCGCCGCGSSCGCSGCCSCSCSC	738
OY	492	GCGGCTTACCACTATGGAACCCAAAGGGGGCCCGCGGTGTCTGTATGGGCGGGGTGCC	551

[illegible]


```

RESULT 1
LOCUS      A87611      1235 bp      DNA
DEFINITION Sequence 8 from Patent WO9836089.
ACCESSION  A87611
VERSION     A87611.1 GI:6736251
KEYWORDS
SOURCE      unidentified.
            unclassified.
ORGANISM    unidentified.
REFERENCE   1 (bases 1 to 1235)
            Flohe, L. and Singh, M.
AUTHORS     TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
TITLE       Patent: WO 9836089-A 20-AUG-1998;
            FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
JOURNAL
FEATURES
            source
            1..1235
            /organism="unidentified"
            /db_xref="taxon:32644"
BASE COUNT  236 a      394 c      385 g      220 t
ORIGIN
Query Match      100.0%; Score 1235; DB 5; Length 1235;
Best Local Similarity 100.0%; Pred. No. 2,7e-168;
Matches 1235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATCTTGCAATTAATGAACTTTCTTCACTGAAAGCGTACAGTATCGAGGGGTAAATC 60
DB      1 ATCTTGCAATTAATGAACTTTCTTCACTGAAAGCGTACAGTATCGAGGGGTAAATC 60
QY      61 ATGCGCGGTGATTCGACCGAGACAAACAAACAAAGATTCGGGTGCGCATCAACCCG 120
DB      61 ATGCGCGGTGATTCGACCGAGACAAACAAACAAAGATTCGGGTGCGCATCAACCCG 120
QY      121 GCGCGCGTGGCGGAACTAACCGGTGCGCATGAGTGTCTATCCAGGCGAGTGGCGGA 180
DB      121 GCGCGCGTGGCGGAACTAACCGGTGCGCATGAGTGTCTATCCAGGCGAGTGGCGGA 180
QY      122 GCGCGCGTGGCGGAACTAACCGGTGCGCATGAGTGTCTATCCAGGCGAGTGGCGGA 180
DB      122 GCGCGCGTGGCGGAACTAACCGGTGCGCATGAGTGTCTATCCAGGCGAGTGGCGGA 180
QY      181 GAGGGGTGGGTATCAACGAGCGGATTTCAAGGGCGGAGGGCGCACTGGTGGCGAC 240
DB      181 GAGGGGTGGGTATCAACGAGCGGATTTCAAGGGCGGAGGGCGCACTGGTGGCGAC 240
QY      241 GCGGACAGGTGAGGCGGACGCGTGAATTTATGCTCAAGGTCAAAAGAACGATAGCGG 300
DB      241 GCGGACAGGTGAGGCGGACGCGTGAATTTATGCTCAAGGTCAAAAGAACGATAGCGG 300
QY      301 GAATAGCGCGCGCTGCGACACGCGGATGTTGTCACGTTCTTGCGCGGCTCAC 360
DB      301 GAATAGCGCGCGCTGCGACACGCGGATGTTGTCACGTTCTTGCGCGGCTCAC 360
QY      361 GTCCTTGACACGATGCTGTTGATTCGGGACGACGATTCGCTTACGAGACCGTCC 420
DB      361 GTCCTTGACACGATGCTGTTGATTCGGGACGACGATTCGCTTACGAGACCGTCC 420
QY      421 AGACGCGCGTAAAGCGGACATCCCTGCTTGCCTGATAGCGAAGTC TCCGCTGCACTG 480
DB      421 AGACGCGCGTAAAGCGGACATCCCTGCTTGCCTGATAGCGAAGTC TCCGCTGCACTG 480
QY      481 CGGCGCAAGTGGCGGTTTACACATGATCGGAACCAAGGGGGCGCGGTGCTGATGG 540
DB      481 CGGCGCAAGTGGCGGTTTACACATGATCGGAACCAAGGGGGCGCGGTGCTGATGG 540
QY      541 GCGGGGTGGCGGCTGCAACCGGCGAGCGTGTGATCGGCGCGGCGGCGGCGGCT 600
DB      541 GCGGGGTGGCGGCTGCAACCGGCGAGCGTGTGATCGGCGCGGCGGCGGCGGCT 600
QY      601 ACAACGCGACCGCGCATGCGCAACGCGATGGGCGGACGCGTTACGCTTACATCAACA 660
DB      601 ACAACGCGACCGCGCATGCGCAACGCGATGGGCGGACGCGTTACGCTTACATCAACA 660
QY      661 TCGACAAACTTGGGCAACTGACGCGGAGTTCTGGCGCGGCTATCGACACTCGCTACTCAT 720

```

```

DB      661 TCGACAAACTTGGGCAACTGACGCGGAGTTCTGGCGCGGCTATCGACACTCGCTACTCAT 720
QY      721 CGGCGTACGAGCTCGAGGGTCCGCTCAAGGTGCGGACCTGCTGATTTGGGCGGCTCTGG 780
DB      721 CGGCGTACGAGCTCGAGGGTCCGCTCAAGGTGCGGACCTGCTGATTTGGGCGGCTCTGG 780
QY      781 TCGCAGGCGCCAAAGGCAACCAATTACTCTGGAATTTACTTTGCGGCAATTTGAAACGAG 840
DB      781 TCGCAGGCGCCAAAGGCAACCAATTACTCTGGAATTTACTTTGCGGCAATTTGAAACGAG 840
QY      841 GTGGGCTACTGGTGAATATACCATTCGACGAGGCGGCGTGTTCGAAGGCTCAGACGCA 900
DB      841 GTGGGCTACTGGTGAATATACCATTCGACGAGGCGGCGTGTTCGAAGGCTCAGACGCA 900
QY      901 CCACCTACGACCCACCCGAGTGGTGGCGGTGACGACGACGCTGTTTACTGCTGGGCAACA 960
DB      901 CCACCTACGACCCACCCGAGTGGTGGCGGTGACGACGACGCTGTTTACTGCTGGGCAACA 960
QY      961 TGCCCGGCTGGTGGCGGAGCGTGCAGCTACGAGCGGTGACCAACGCGCATGCGGCTATG 1020
DB      961 TGCCCGGCTGGTGGCGGAGCGTGCAGCTACGAGCGGTGACCAACGCGCATGCGGCTATG 1020
QY      1021 TGCTGACCTTGGCGGACCATGCTGCGGCGGCGGCGTCCGCTCGAATCCGCGCATACCA 1080
DB      1021 TGCTGACCTTGGCGGACCATGCTGCGGCGGCGGCGTCCGCTCGAATCCGCGCATACCA 1080
QY      1081 AAGGCTTTTCGACGACGACGAGAGGCGGCTTACTGTCCGAAGGCGGTGGCCACGACTGGGG 1140
DB      1081 AAGGCTTTTCGACGACGACGAGAGGCGGCTTACTGTCCGAAGGCGGTGGCCACGACTGGGG 1140
QY      1141 TGCCG TCGACGAGCGGCGGCGGCGTGGCTGAGCTGCTGCGCGCTGCTTACCGCAGC 1200
DB      1141 TGCCGTTACCGAGCGGCGGCGGCGTGGCTGAGCTGCTGCGCGCTGCTTACCGCAGC 1200
QY      1201 ACAGCTCGGAGTAAAGGAGCGATGATGTGCGCC 1235
DB      1201 ACAGCTCGGAGTAAAGGAGCGATGATGTGCGCC 1235

RESULT 2
LOCUS      A89751      1235 bp      DNA
DEFINITION Sequence 8 from Patent WO9832862.
ACCESSION  A89751
VERSION     A89751.1 GI:6738285
KEYWORDS
SOURCE      unidentified.
            unclassified.
ORGANISM    unidentified.
REFERENCE   1 (bases 1 to 1235)
            Flohe, L. and Singh, M.
AUTHORS     L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
TITLE       Patent: WO 9832862-A 30-JUL-1998;
            FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
JOURNAL
FEATURES
            source
            1..1235
            /organism="unidentified"
            /db_xref="taxon:32644"
BASE COUNT  236 a      394 c      385 g      220 t
ORIGIN
Query Match      100.0%; Score 1235; DB 5; Length 1235;
Best Local Similarity 100.0%; Pred. No. 2,7e-168;
Matches 1235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATCTTGCAATTAATGAACTTTCTTCACTGAAAGCGTACAGTATCGAGGGGTAAATC 60
DB      1 ATCTTGCAATTAATGAACTTTCTTCACTGAAAGCGTACAGTATCGAGGGGTAAATC 60
QY      61 ATGCGCGGTGATTCGACCGAGACAAACAAACAAAGATTCGGGTGCGCATCAACCCG 120

```

D	b	61	ATGCGCGTGGGTATTCCGACCGAGAACCAAAAACAAAGAAATTCGGGGTGGCCATCACCCG	120
O	y	121	GCGGCGTGCGGGACTAACCCGCTGCGCATAGAGTGTCTCATCCAGCGAGTCCGGA	180
D	b	121	GCGGCGTGCGGGAAGTAACCCGCTGCGCCATAGAGTGTCTCATCCAGCGAGTCCGGA	180
O	y	181	GAGGGCTGGGTATCCAGGACGGGAAATTTAAAGGCGGAGGGCGCAATGGTGGCGAC	240
D	b	181	GAGGGCTGGGTATCCAGGACGGGAAATTTAAAGGCGGAGGGCGCAATGGTGGCGAC	240
O	y	241	GCCGCAAGGTGTGGGCGCAGACGCTGATTTATGCTCAAGGTCAAAAACCGATAGCGCG	300
D	b	241	GCCGCAAGGTGTGGGCGCAGACGCTGATTTATGCTCAAGGTCAAAAACCGATAGCGCG	300
O	y	301	GAATACGGGCCCCCTGGCAGACAGGGGCGATCTGTTCACGTCTTGGCATTTGGCGGGGTAC	360
D	b	301	GAATACGGGCCCCCTGGCAGACAGGGGCGATCTGTTCACGTCTTGGCATTTGGCGGGGTAC	360
O	y	361	GTCCTTGACAGCATGAGCTGTGTGGATTCGGGACCAAGTCATATGGCTACAGACCGTC	420
D	b	361	GTCCTTGACAGCATGAGCTGTGTGGATTCGGGACCAAGTCATATGGCTACAGACCGTC	420
O	y	421	AGACCGCCGAAGGCGCACTACCCCTGCTTGCCTCCCGATGACGGAATGCGCGGTGCATCG	480
D	b	421	AGACCGCCGAAGGCGCACTACCCCTGCTTGCCTCCCGATGACGGAATGCGCGGTGCATCG	480
O	y	481	CGGCGCAGATGGGGGCTTACACCTGATGGGAACCAAGGGGGCGCGGTGTGATAGG	540
D	b	481	CGGCGCAGATGGGGGCTTACACCTGATGGGAACCAAGGGGGCGCGGTGTGATAGG	540
O	y	541	GCGGGGTGCCCCGCGCTGCAACCGGCGCAGCTCGTGTGATCGGCGCGGCAACCGCGCT	600
D	b	541	GCGGGGTGCCCCGCGCTGCAACCGGCGCAGCTCGTGTGATCGGCGCGGCAACCGCGCT	600
O	y	601	ACAAGGAGCCCCGATATGCCAAGGGCATGGGGCGGACCGTTACGGTCTAAGCATCAACA	660
D	b	601	ACAAGGAGCCCCGATATGCCAAGGGCATGGGGCGGACCGTTACGGTCTAAGCATCAACA	660
O	y	661	TGCAACAACTTCGGCACTGCGACGCGCAGATTCGCGGCGGATCCCACTGCTCCTACTAT	720
D	b	661	TGCAACAACTTCGGCACTGCGACGCGCAGATTCGCGGCGGATCCCACTGCTCCTACTAT	720
O	y	721	GCGGCTACGAGCTCGAAGGTGCGTCAACGTCGCGACCTGGATGGGGCGGCTCTGG	780
D	b	721	GCGGCTACGAGCTCGAAGGTGCGTCAACGTCGCGACCTGGATGGGGCGGCTCTGG	780
O	y	781	TGCCAGGGGCCAAGGCAACCCAAATATGTCGAAATCACTGTGCGGCAATAGGAACAG	840
D	b	781	TGCCAGGGGCCAAGGCAACCCAAATATGTCGAAATCACTGTGCGGCAATAGGAACAG	840
O	y	841	GTCGGTACTGTGTGATATAGCCATGACACAGGGCGGCTGTTTCGAAGGCTACACAGCA	900
D	b	841	GTCGGTACTGTGTGATATAGCCATGACACAGGGCGGCTGTTTCGAAGGCTACACAGCA	900
O	y	901	CCACCTACGACACACCGACGTTGGCGGTGACAGACACGTTTATACGTGCGGGGACACA	960
D	b	901	CCACCTACGACACACCGACGTTGGCGGTGACAGACACGTTTATACGTGCGGGGACACA	960
O	y	961	TGCCCGGCTCGGTGCGAAGACGTGCACTACGCGGTGACCAACGGGAGATGCGGTATG	1020
D	b	961	TGCCCGGCTCGGTGCGAAGACGTGCACTACGCGGTGACCAACGGGAGATGCGGTATG	1020
O	y	1021	TGCTCGAGCTTGGCGACATAGGGTGGGGGGCGCGGCGGTGCAATCCGGCATAGACA	1080
D	b	1021	TGCTCGAGCTTGGCGACATAGGGTGGGGGGCGCGGCGGTGCAATCCGGCATAGACA	1080
O	y	1081	AAGGTCTTTCGACGACAGAAAGGGCGTTACTGTCCGAACGGGTGGGCAACGACCTGGGGG	1140
D	b	1081	AAGGTCTTTCGACGACAGAAAGGGCGTTACTGTCCGAACGGGTGGGCAACGACCTGGGGG	1140
O	y	1141	TGCGGTTACGAGACCCGCGCAAGGCTGTGGCTTCACTTCGCGCGTGTATACGCCAGC	1200
D	b	1141	TGCGGTTACGAGACCCGCGCAAGGCTGTGGCTTCACTTCGCGCGTGTATACGCCAGC	1200

QY	1201	ACAGTCGGGAGTAAGGAAGAAGCATGTGC	1235
Db	1201	ACAGTCGGGAGTAAGGAAGAAGCATGTGC	1235
RESULT	3		
LOCUS	A87609	A87609	1235 bp DNA
DEFINITION		Sequence 6 from Patent WO9836089.	
ACCESSION		A87609	
VERSION		A87609.1	GI:6736249
KEYWORDS		.	
SOURCE		unidentified.	
ORGANISM		unidentified	
		unclassified.	
REFERENCE		1 (bases 1 to 1235)	
AUTHORS		Flohe,L. and Singh,M.	
TITLE		TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE	
JOURNAL		Patent: WO 9836089-A-20-AUG-1998;	
FEATURES		FLHGE LEOPOLD (DE); SINGH MAHAVIR (DE)	
Source		location/Qualifiers	
		1..1235	
		/organism="unidentified"	
		/db_xref="taxon:32644"	
BASE COUNT	235 a	395 c	384 g 220 t
ORIGIN			1 others

Query Match	99.8%;	Score 1232.4;	DB 5;	Length 1235;
Best Local Similarity	99.8%;	Pred. No. 6.3e-168;		
Matches 1233;	Conservative	0;	Mismatches 2;	Indels 0;
				Gaps 0;

QY	1	ATTGTGAGATTAAATCGAACTTTTCTTCACATGAAAGCTACAGATATGAGAGGGGTATC	60
Db	1	ATCTTTCGAGATTAAATCGAACTTTTCTTCACATGAAAGCTACAGATATGAGAGGGGTATC	60
QY	61	ATGCGCGTCGGTATTCCGACCGAGACGCAAAAACAACGAATTCGGGTGGCCATACCCCG	120
Db	61	ATGCGCGTCGGTATTCCGACCGAGACGCAAAAACAACGAATTCGGGTGGCCATACCCCG	120
QY	121	GCCGGCGTCGCGGAATTAACCGTCGTGGCCATGAGTGGCTCATCCAGGACGAGTGGCCGA	180
Db	121	GCCGGCGTCGCGGAACCTAACCGTCGTGGCCATGAGTGGCTCATCCAGGACGAGTGGCCGA	180
QY	181	GAGGCGTCGGCTATCCACCGACCGGATTTCAAGCGCGGACGCGCACTGTGGGACC	240
Db	181	GAGGCGTCGGGTATCCACCGACCGGATTTCAAGCGCGGACGCGCACTGTGGGACC	240
QY	241	GCCGACGAGGTGTGGGGCCGACGGCTGATTTATTGGTCAAGGTCAAAAGAACGATAGCGGG	300
Db	241	GCCGACGAGGTGTGGGGCCGACGGCTGATTTATTGGTCAAGGTCAAAAGAACGATAGCGGG	300
QY	301	GAAATACGAGCCGCTCCGACACGAGCGCATCTTGTACAGTTTTCGATTGGCCGCGTAC	360
Db	301	GAAATACGAGCCGCTCCGACACGAGCGCATCTTGTACAGTTTTCGATTGGCCGCGTAC	360
QY	361	GTGCTTGCAACCGAGTGGTGTGTGGATTCCGGCCACGACGTCAATTGGCTACGAGACGCTCC	420
Db	361	GTGCTTGCAACCGATCCGTGTGTGGATTCCGGCCACGACGTCAATTGGCTACGAGACGCTCC	420
QY	421	AGACCGCCGAAGGCGCATATACCCCGTGTGGCCCGGATGAGGAAGTGGCGGTGTGACTCG	480
Db	421	AGACCGCCGACGAGCGCATATACCCCGTGTGGCCCGGATGAGGAAGTGGCGGTGTGACTCG	480
QY	481	CCGCGCAGGTTGGCGCTTACCACTGATGCGAACCCAAAGGGGGCGCGGTGTGCTGATGG	540
Db	481	CCGCGCAGGTTGGCGCTTACCACTGATGCGAACCCAAAGGGGGCGCGGTGTGCTGATGG	540
QY	541	GCGGGGTCGCCGGCTCTAAACCGGCGAGCTGTGTGTATCGGCGCGGACACGCGCGGCT	600
Db	541	GCGGGGTCGCCGGCTCTAAACCGGCGAGCTGTGTGTATCTCGCGCGGACACGCGCGGCT	600

QY 601 ACAACGACCCGATCGCCAAACGAGCATGGGCGGACCGTTACGCTTACATCAACA 660
DB 601 ACAACGACCCGATCGCCAAACGAGCATGGGCGGACCGTTACGCTTACATCAACA 660
QY 661 TCGACAACCTTGGGCACTCGACGCGGAGTTCTGGGCGGATCCACTCGCTACTCAT 720
DB 661 TCGACAACCTTGGGCACTCGACGCGGAGTTCTGGGCGGATCCACTCGCTACTCAT 720
QY 721 CGGCGTACGAGCTCGAGGGTGGCGGCTCAACGTCGACCTGGGATGGGCGGCTCTGG 780
DB 721 CGGCGTACGAGCTCGAGGGTGGCGGCTCAACGTCGACCTGGGATGGGCGGCTCTGG 780
QY 781 TCGCAGGCGCCCAAGGACCCCAATTAAGTCTCAATTCAGTTGCGGATATGAACAG 840
DB 781 TCGCAGGCGCCCAAGGACCCCAATTAAGTCTCAATTCAGTTGCGGATATGAACAG 840
QY 841 GTGCGGTACTGGTGGATATAGCATGACACAGGCGGCTGTTTCAAGGCTCAGACGGA 900
DB 841 GTGCGGTACTGGTGGATATAGCATGACACAGGCGGCTGTTTCAAGGCTCAGACGGA 900
QY 901 CCACCTACGACGACCCGACGTTGCGGCTGACAGACGCTGTTTACTGCGTGGGAGCA 960
DB 901 CCACCTACGACGACCCGACGTTGCGGCTGACAGACGCTGTTTACTGCGTGGGAGCA 960
QY 961 TCGCCGCTCGGTGGCGGAAGACGTCGACCTAGCGGCTGACCAACGCGAGATGCCGTATG 1020
DB 961 TCGCCGCTCGGTGGCGGAAGACGTCGACCTAGCGGCTGACCAACGCGAGATGCCGTATG 1020
QY 1021 TCGCTGAGCTTGGCGGACATGGTGGCGGCGGCGGCTGGGATCCGGCATCCGCACTAGCCA 1080
DB 1021 TCGCTGAGCTTGGCGGACATGGTGGCGGCGGCGGCTGGGATCCGGCATCCGCACTAGCCA 1080
QY 1081 AAGGCTTTTCAGACGACGAAAGGCGGCTTACTGTCGACGAGGCGGACGACGCTGGGG 1140
DB 1081 AAGGCTTTTCAGACGACGAAAGGCGGCTTACTGTCGACGAGGCGGACGACGCTGGGG 1140
QY 1141 TCGCGCTTACCGACGCGGCGGACGCTGCTGCTGACTCTCGGCGGCTGTTACGCCGAGC 1200
DB 1141 TCGCGCTTACCGACGCGGCGGACGCTGCTGCTGACTCTCGGCGGCTGTTACGCCGAGC 1200
QY 1201 ACACGTCGGGAGTAAAGGAGGAGGATGTCGGGCG 1235
DB 1201 ACACGTCGGGAGTAAAGGAGGAGGATGTCGGGCG 1235

RESULT 4
LOCUS AB9749 1235 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 6 from Patent WO9832862.
ACCESSION AB9749
VERSION AB9749.1 GI:6738283
KEYWORDS
SOURCE .
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1235)
AUTHORS Flohe, U. and Singh, M.
TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
JOURNAL Patent: WO 9832862-A 30-JUL-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
SOURCE 1. 1235
Location/Qualifiers
/db_xref="taxon:32644"
BASE COUNT 235 a 395 c 384 g 220 t 1 others
ORIGIN

Query Match 99.8%; Score 1232.4; DB 5; Length 1235;
Best Local Similarity 99.8%; Pred. No. 6.3e-168;
Matches 1233; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTTGCAGATTAATGCAACTTCTTCACACGTAAGCGTACAGTATCGAGAGGGTAATC 60

DB 1 ACTTGCAGATTAATGCAACTTCTTCACACGTAAGCGTACAGTATCGAGAGGGTAATC 60
QY 61 ATGCGCTGCGGTATTCGACCGAGACCAAAACAGATTCCTGGGTGGCATACCCCG 120
DB 61 ATGCGCTGCGGTATTCGACCGAGACCAAAACAGATTCCTGGGTGGCATACCCCG 120
QY 121 GCGGCGTCCGGGAATCAACCCCTGCTGGCATAGAGTGTCTATCCAGGAGGTCCGGA 180
DB 121 GCGGCGTCCGGGAATCAACCCCTGCTGGCATAGAGTGTCTATCCAGGAGGTCCGGA 180
QY 181 GAGGCGTCCGGATATACCGAGCGGATTTCAAGGGGCGAGGCGGCACTGCTGGGACC 240
DB 181 GAGGCGTCCGGATATACCGAGCGGATTTCAAGGGGCGAGGCGGCACTGCTGGGACC 240
QY 241 GCGGACCAAGTGTGGGCGGACGCTGATTTATTTCTCAAGGTCAAGAACGATAGCGCG 300
DB 241 GCGGACCAAGTGTGGGCGGACGCTGATTTATTTCTCAAGGTCAAGAACGATAGCGCG 300
QY 301 GAATACGGCGCGCTGCGACGAGGGGATGCTGTACCTTCTTGATTTGGCGCGCTAC 360
DB 301 GAATACGGCGCGCTGCGACGAGGGGATGCTGTACCTTCTTGATTTGGCGCGCTAC 360
QY 361 GTGCTTGCACGATGCGCTTGTGGATTCGGGACACGCTCAATTTGCTACGAGACCGTCC 420
DB 361 GTGCTTGCACGATGCGCTTGTGGATTCGGGACACGCTCAATTTGCTACGAGACCGTCC 420
QY 421 AGACCGCGGAAAGCGGACATACCCCTGCTGGCGGATGAGGAGTCCGCGGTGACTCG 480
DB 421 AGACCGCGGAAAGCGGACATACCCCTGCTGGCGGATGAGGAGTCCGCGGTGACTCG 480
QY 481 CCGCCCAAGTTGGCGCTTACCACTGATGACGACCAAGGGGCGGCTGTGCTGATGG 540
DB 481 CCGCCCAAGTTGGCGCTTACCACTGATGACGACCAAGGGGCGGCTGTGCTGATGG 540
QY 541 GCGGCGTCCCGGCGCT...ACGCGCGACGCTGCTGTGATGAC GCGCGGACCGCGGCT 600
DB 541 GCGGCGTCCCGGCGCT...ACGCGCGACGCTGCTGTGATGACGCGCGGCGCT 600
QY 601 ACAACGACCCGCGATTCGCCCAAGGAGGAGGCGGCGGCTTACGCTTACATCAACA 660
DB 601 ACAACGACCCGCGATTCGCCCAAGGAGGAGGCGGCGGCTTACGCTTACATCAACA 660
QY 661 TCGACAACCTTGGGCACTCGACGCGGAGTTCTGGGCGGATCCACTCGCTACTCAT 720
DB 661 TCGACAACCTTGGGCACTCGACGCGGAGTTCTGGGCGGATCCACTCGCTACTCAT 720
QY 721 CGGCGTACGAGCTCGAGGGTGGCGGCTCAACGTCGACCTGGGATGGGCGGCTCTGG 780
DB 721 CGGCGTACGAGCTCGAGGGTGGCGGCTCAACGTCGACCTGGGATGGGCGGCTCTGG 780
QY 781 TCGCAGGCGCCCAAGGACCCCAATTAAGTCTCAATTCAGTTGCGGATATGAACAG 840
DB 781 TCGCAGGCGCCCAAGGACCCCAATTAAGTCTCAATTCAGTTGCGGATATGAACAG 840
QY 841 GTGCGGTACTGGTGGATATAGCATGACACAGGCGGCTGTTTCAAGGCTCAGACGGA 900
DB 841 GTGCGGTACTGGTGGATATAGCATGACACAGGCGGCTGTTTCAAGGCTCAGACGGA 900
QY 901 CCACCTACGACGACCCGACGTTGCGGCTGACAGACGCTGTTTACTGCGTGGGAGCA 960
DB 901 CCACCTACGACGACCCGACGTTGCGGCTGACAGACGCTGTTTACTGCGTGGGAGCA 960
QY 961 TCGCCGCTCGGTGGCGGAAGACGTCGACCTAGCGGCTGACCAACGCGAGATGCCGTATG 1020
DB 961 TCGCCGCTCGGTGGCGGAAGACGTCGACCTAGCGGCTGACCAACGCGAGATGCCGTATG 1020
QY 1021 TCGCTGAGCTTGGCGGACATGGTGGCGGCGGCGGCTGGGATCCGGCATCCGCACTAGCCA 1080
DB 1021 TCGCTGAGCTTGGCGGACATGGTGGCGGCGGCGGCTGGGATCCGGCATCCGCACTAGCCA 1080
QY 1081 AAGGCTTTTCAGACGACGAAAGGCGGCTTACTGTCGAAAGGCTGGGCGACGACCTGGGG 1140

Db 1081 AAGGTTTCGACGACGAGGAGGCGCTACTGTCCGACGAGGCGGCGCCACGACTGCGGG 1140

QY 1141 TGCCGTTTCAACGAGCCCGCAGCGTGTGCTGACTCTGCGCGCTGTATACCGGACG 1200

Db 1141 TGCCGTTTCAACGAGCCCGCAGCGTGTGCTGACTCTGCGCGCTGTATACCGGACG 1200

QY 1201 ACACGTCGAGAGTAAGGAAGCGATGATGTCGGCC 1235

Db 1201 ACACGTCGAGAGTAAGGAAGCGATGATGTCGGCC 1235

RESULT 5

LOCUS A87610 1229 bp DNA PAT 22-JAN-2000

DEFINITION Sequence 7 from Patent WO9836089.

ACCESSION A87610

VERSION A87610.1 GI:6736250

KEYWORDS

SOURCE unidentified.

ORGANISM unidentified.

REFERENCE 1 (bases 1 to 1229)

AUTHORS Flohe, L. and Singh, M.

TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE

JOURNAL Patent: WO 9836089-A-20-AUG-1998;

FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)

FEATURES

source 1..1229

/organism="unidentified"

/db_xref="taxon:32644"

BASE COUNT 236 a 391 c 382 g 219 t 1 others

ORIGIN

Query Match 99.48; Score 1228; DB 5; Length 1229;

Best Local Similarity 99.98; Freq. No. 2.7e-167;

Matches 1228; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATCTGAGATTATCGACTTTCTTCACTGAGAGTACAGTATCGAGGGGTATC 60

Db 1 ATCTGAGATTATCGACTTTCTTCACTGAGAGTACAGTATCGAGGGGTATC 60

QY 61 ATGCGCTGCTGATTCGACGAGAGCCAAACACGAATTCGGGTGCGCATCACC 120

Db 61 ATGCGCTGCTGATTCGACGAGAGCCAAACACGAATTCGGGTGCGCATCACC 120

QY 121 GCGGCGCTGCGGAGTAACCCGTCGTGGCATAGGTGCTCATCCAGGAGTCCGGA 180

Db 121 GCGGCGCTGCGGAGTAACCCGTCGTGGCATAGGTGCTCATCCAGGAGTCCGGA 180

QY 181 GAGGCTCGCTATACGACGAGCGGATTCAGAGCGGCGGCGCAACTGTCGAC 240

Db 181 GAGGCTCGCTATACGACGAGCGGATTCAGAGCGGCGGCGCAACTGTCGAC 240

QY 241 GCGGACGAGTGTGGCGGAGCGTGTATTCGTCAGAGTCAGAGACGATGCGCG 300

Db 241 GCGGACGAGTGTGGCGGAGCGTGTATTCGTCAGAGTCAGAGACGATGCGCG 300

QY 301 GAATACGAGCGCTGCGACAGCGGAGTCTGTTCACGTTCTTGCAATTTGGCGCGTAC 360

Db 301 GAATACGAGCGCTGCGACAGCGGAGTCTGTTCACGTTCTTGCAATTTGGCGCGTAC 360

QY 361 GTGCTTGCACGAGTGTGTGTGATTCGCGGACGACGATGCTACGAGACGCTCC 420

Db 361 GTGCTTGCACGAGTGTGTGTGATTCGCGGACGACGATGCTACGAGACGCTCC 420

QY 421 AGACGCGGAGGCGCACTACCTGCTTGGCCGATGAGGAGTGC CGGTGACTCG 480

Db 421 AGACGCGGAGGCGCACTACCTGCTTGGCCGATGAGGAGTGC CGGTGACTCG 480

QY 481 CCGCCAGAGTGTGGGCTTACCACTGATGCAACCAAGGGGCGCGGTGTGATGG 540

Db 481 CCGCCAGAGTGTGGGCTTACCACTGATGCAACCAAGGGGCGCGGTGTGATGG 540

QY 541 GCGGGGTGCGCGCGCTCGAACCGCGGACGTCGTGTGATGCGGCGCGGCGCGGCT 600

Db 541 GCGGGGTGCGCGCGCTCGAACCGCGGACGTCGTGTGATGCGGCGCGGCGCGGCT 600

QY 601 ACACGAGCGCCGCACTCCCAACGAGCATGAGGCGGCGGCGGTTACGTTTACATCAACA 660

Db 601 ACACGAGCGCCGCACTCCCAACGAGCATGAGGCGGCGGCGGTTACGTTTACATCAACA 660

QY 661 TCGACAAACTTGGCACTCGACGCGGAGTTCTGCGCGCGATCCACTGCTACTCAT 720

Db 661 TCGACAAACTTGGCACTCGACGCGGAGTTCTGCGCGCGATCCACTGCTACTCAT 720

QY 721 CGGCTACGAGCTGAG, TGCCGTCAAACGTCGCGACCTGATGTTGGGCGGCTG 780

Db 721 CGGCTACGAGCTGAGGAGTTCGCGCAACGTCGCGACCTGATGTTGGGCGGCTG 780

QY 781 TCGCAGGCGCCCAAGCCACCAATTAGTCTCGAATTCATTGTCGCAATGAAACGAG 840

Db 781 TCGCAGGCGCCCAAGCCACCAATTAGTCTCGAATTCATTGTCGCAATGAAACGAG 840

QY 841 TCGGCTACTGCTGATATAGCATCGACGAGGCGGCTGTTGGAAGCTCAGACCGA 900

Db 841 TCGGCTACTGCTGATATAGCATCGACGAGGCGGCTGTTGGAAGCTCAGACCGA 900

QY 901 CCACCTACGACCCGACGAGTTCGCGCTGACAGACGCTGTTTACTGCTGCGGGAACA 960

Db 901 CCACCTACGACCCGACGAGTTCGCGCTGACAGACGCTGTTTACTGCTGCGGGAACA 960

QY 961 TCGCGGCTGCTGCGGAGAGAGTGCAGTACGCGTACGACGAGGAGTCCGTTATG 1020

Db 961 TCGCGGCTGCTGCGGAGAGTGCAGTACGCGTACGACGAGGAGTCCGTTATG 1020

QY 1021 TCGTCAAGCTTCCGCACTGAGTGTGCGGCGGCGGCTGCAATCGGACCTAGCA 1080

Db 1021 TCGTCAAGCTTCCGCACTGAGTGTGCGGCGGCGGCTGCAATCGGACCTAGCA 1080

QY 1081 AAGGTTTTCGACGACGAGGAGGCGTGTCTGTCGACGAGTGTGCGGCGGAGG 1140

Db 1081 AAGGTTTTCGACGACGAGGAGGCGTGTCTGTCGACGAGTGTGCGGCGGAGG 1140

QY 1141 TGCCGTTTCAACGAGCCCGCAGCGTGTGCTGACTCGGCGGCTGTGAGCGGAGC 1200

Db 1141 TGCCGTTTCAACGAGCCCGCAGCGTGTGCTGACTCGGCGGCTGTGAGCGGAGC 1200

QY 1201 ACACGTCGAGAGTAAGGAAGCGATGATG 1229

Db 1201 ACACGTCGAGAGTAAGGAAGCGATGATG 1229

RESULT 6

LOCUS A89750 1229 bp DNA PAT 22-JAN-2000

DEFINITION Sequence 7 from Patent WO9832862.

ACCESSION A89750

VERSION A89750.1 GI:6738284

KEYWORDS

SOURCE unidentified.

ORGANISM unidentified.

REFERENCE 1 (bases 1 to 1229)

AUTHORS Flohe, L. and Singh, M.

TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM

JOURNAL Patent: WO 9832862-A-30-JUN-1998;

FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)

FEATURES

source 1..1229

/organism="unidentified"

/db_xref="taxon:32644"

BASE COUNT 236 a 391 c 382 g 219 t 1 others

ORIGIN

Db 421 CAGACCCGCGACGCGACACTACCCCTGCTGCCCCGATGAGCAAGTCCGCGGTGACTC 480
 QY 480 GCGGCCAGTTGGCGCTTACACACTGATGCCAAGGAGGGGCGCGGTGCTGATG 539
 Db 481 GCGGCCAGTTGGCGCTTACACACTGATGCCAAGGAGGGGCGCGGTGCTGATG 540
 QY 540 GCGGCCAGTTGGCGCTTACACACTGATGCCAAGGAGGGGCGCGGTGCTGATG 599
 Db 541 GCGGCCAGTTGGCGCTTACACACTGATGCCAAGGAGGGGCGCGGTGCTGATG 600
 QY 600 TACACGAGCCCGCATCCGCAACGGCATGCGCGCAAGCTTACGGTTCTAGACATCAAC 659
 Db 601 TACACGAGCCCGCATCCGCAACGGCATGCGCGCAAGCTTACGGTTCTAGACATCAAC 660
 QY 660 ATGACAACTTGGGACATCGACGCGGATGCTGCGCGCGGATCCACATCGCTACTCA 719
 Db 661 ATGACAACTTGGGACATCGACGCGGATGCTGCGCGCGGATCCACATCGCTACTCA 720
 QY 720 TCGGCTTACGAGCTGAGGGTGGCTGCAACGTTGCGGACCTGATGATGCGCGCTG 779
 Db 721 TCGGCTTACGAGCTGAGGGTGGCTGCAACGTTGCGGACCTGATGATGCGCGCTG 780
 QY 780 GTGCCAGGGCGCAAGGACCCCAATAGTCTGCAATTCCTGCTGCGCATATGAACA 839
 Db 781 GTGCCAGGGCGCAAGGACCCCAATAGTCTGCAATTCCTGCTGCGCATATGAACA 840
 QY 840 GGTGCGGTACTGATGATATAGCCATCGACGAGCGGCTGTTTCGAAGGCTCAGCAGCG 899
 Db 841 GGTGCGGTACTGATGATATAGCCATCGACGAGCGGCTGTTTCGAAGGCTCAGCAGCG 900
 QY 900 ACCACTAGACACACCGGACGCTGCGCGTGCACGACAGCTGTTTACTGCTGCGGAAAC 959
 Db 901 ACCACTAGACACACCGGACGCTGCGCGTGCACGACAGCTGTTTACTGCTGCGGAAAC 960
 QY 960 ATGCCCGCTCGGTGCGCAAGAGCTGACCTGACGCTGACCAAGCGAGATGCCGTAT 1019
 Db 961 ATGCCCGCTCGGTGCGCAAGAGCTGACCTGACGCTGACCAAGCGAGATGCCGTAT 1020
 QY 1020 GTGCTGAGCTTGGCGACCATGCTGCGGCGGCGGCTGCGGCTGCAATCGGACATAGCC 1079
 Db 1021 GTGCTGAGCTTGGCGACCATGCTGCGGCGGCGGCTGCGGCTGCAATCGGACATAGCC 1080
 QY 1080 AAGGCTTTCGACGCGACGAGAGGGGGCTTACTGTCGAGAGGGTGGCCACCGACCTGGGG 1139
 Db 1081 AAGGCTTTCGACGCGACGAGAGGGGGCTTACTGTCGAGAGGGTGGCCACCGACCTGGGG 1140
 QY 1140 GTGCGCTTACGAGCGCGCGACGCTGCTGCTGACTCTGCGCGCTGCTTACGCGGAG 1199
 Db 1141 GTGCGCTTACGAGCGCGCGACGCTGCTGCTGACTCTGCGCGCTGCTTACGCGGAG 1200
 QY 1200 CACACTCGGAGTAGAGGAGGATGATGCTGCGCC 1235
 Db 1201 CACACTCGGAGTAGAGGAGGATGATGCTGCGCC 1236

RESULT 8
 LOCUS A89753 1236 bp DNA
 DEFINITION A89753 Sequence 10 from Patent WO9832862. PAT 22-JAN-2000
 VERSION A89753.1 GI:6738287
 KEYWORDS
 ORGANISM
 SOURCE unidentified.
 unclassified.
 unclassified.
 unclassified.
 REFERENCES 1 (bases 1 to 1236)
 FLOHE, L. and Singh, M.
 TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
 JOURNAL Patent: WO 9832862-A 30-JUL-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE);
 FEATURES
 source 1..1236
 /organism="unidentified"

BASE COUNT 236 a 395 c 385 g 220 t
 ORIGIN
 Query Match 99.0%; Score 1222.4; DB 5; Length 1236;
 Best local similarity 99.8%; Pred. No. 1.7e-166;
 Matches 1234; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

/db_xref="taxon:32644"
 Db 1 ATCTTGCAGATTATATCGAATCTTCTTACACATGAAAGCGTACAGTATCGAGGGGTATC 60
 Db 1 ATCTTGCAGATTATATCGAATCTTCTTACACATGAAAGCGTACAGTATCGAGGGGTATC 60
 QY 61 ATGCGCGGTGGTATTCGCGCGAGCAAAACCAAGATTCGGGGTGGCATCACCCG 120
 Db 61 ATGCGCGGTGGTATTCGCGCGAGCAAAACCAAGATTCGGGGTGGCATCACCCG 120
 QY 121 GCGGGGTGCGGAACTAACCCGCTGCGCATGAGGTGCTATCCAGGAGGTCGCGGA 180
 Db 121 GCGGGGTGCGGAACTAACCCGCTGCGCATGAGGTGCTATCCAGGAGGTCGCGGA 180
 QY 181 GAGGGTGGGCTATACCGACGCGGATTTCAAGCGCGAGCGCGCAACTGCTGGCACC 240
 Db 181 GAGGGTGGGCTATACCGACGCGGATTTCAAGCGCGAGCGCGCAACTGCTGGCACC 240
 QY 241 GCCGACGAGGTGGGCGGAGCGCTGATTTATGTCMAAGTCAAGAAACCGATAGCGGG 300
 Db 241 GCCGACGAGGTGGGCGGAGCGCTGATTTATGTCMAAGTCAAGAAACCGATAGCGGG 300
 QY 301 GAAT..CGGCGGCTCGGACGACGAGGCG..GATCTTGTTCACGTTCTTGCATTGGCCGCTCA 359
 Db 301 GAAT..CGGCGGCTCGGACGACGAGGCG..GATCTTGTTCACGTTCTTGCATTGGCCGCTCA 360
 QY 360 CGTGTTCGACGAGTGGCTGTTGATTCGCGACACGCTCAATTCCTACGACACCGTTC 419
 Db 361 CGTGTTCGACGAGTGGCTGTTGATTCGCGACACGCTCAATTCCTACGACACCGTTC 420
 QY 420 CAGACCGCGAGAGGCGACATACCTGCTGCGCGGATGAGGAAGTGGCGGTGAGCTC 479
 Db 421 CAGACCGCGAGAGGCGACATACCTGCTGCGCGGATGAGGAAGTGGCGGTGAGCTC 480
 QY 480 GCGCGCGAGTGGCGCTTACCACTGATCGAACCAAGGGGCGCGGTGCTGATG 539
 Db 481 GCGCGCGAGTGGCGCTTACCACTGATCGAACCAAGGGGCGCGGTGCTGATG 540
 QY 540 GCGGGGTGCGCGGCTGCAACCGGCGGACGCTGCTGATGCGGCGCGGCGCGG 599
 Db 541 GCGGGGTGCGCGGCTGCAACCGGCGGACGCTGCTGATGCGGCGCGGCGCGG 600
 QY 600 TACACGAGCCCGCATCCGCAACGGCATGCGCGGACGCTTACGAGTTCATACATCAAC 659
 Db 601 TACACGAGCCCGCATCCGCAACGGCATGCGCGGACGCTTACGAGTTCATACATCAAC 660
 QY 660 ATGACAACTTGGGACATCGACGCGGATGCTGCGCGCGGATCCACATCGCTACTCA 719
 Db 661 ATGACAACTTGGGACATCGACGCGGATGCTGCGCGCGGATCCACATCGCTACTCA 720
 QY 720 TCGGCTTACGAGCTGAGGGTGGCTGCAACGTTGCGGACCTGATGATGCGCGCTG 779
 Db 721 TCGGCTTACGAGCTGAGGGTGGCTGCAACGTTGCGGACCTGATGATGCGCGCTG 780
 QY 780 GTGCCAGGGCGCAAGGACCCCAATAGTCTGCAATTCCTGCTGCGCATATGAACA 839
 Db 781 GTGCCAGGGCGCAAGGACCCCAATAGTCTGCAATTCCTGCTGCGCATATGAACA 840
 QY 840 GGTGCGGTACTGATGATATAGCCATCGACGAGCGGCTGTTTCGAAGGCTCAGCAGCG 899
 Db 841 GGTGCGGTACTGATGATATAGCCATCGACGAGCGGCTGTTTCGAAGGCTCAGCAGCG 900
 QY 900 ACCACTAGACACACCGGACGCTGCGCGTGCACGACAGCTGTTTACTGCTGCGGAAAC 959
 Db 901 ACCACTAGACACACCGGACGCTGCGCGTGCACGACAGCTGTTTACTGCTGCGGAAAC 960

QY	1020	GTGCTCAGAGCTTCCACACATGGCTGAGCGGCGGCTGCCGCTGCATATCCGGACATAAGC	1079
Db	1021	GTGCTCAGAGCTTCCCGACATATGGCTGCGGGCGCGCTGCCGGTTCATATCCGGACATAAGC	1080
QY	1080	AAAGGTCTTTCGAGCGACGAGAAAGGGGCGTTACTCTCGGAAGGGGTGGCCACCGACACTGGGG	1139
Db	1081	AAAGGTCTTTCGAGCGACGAGAAAGGGGCGTTACTCTCGGAAGGGGTGGCCACCGACACTGGGG	1140
QY	1140	GTGCGCTTCACCGAGCCCCCGCCAGCGTGTCTGAGCTCTGGCGCGCTGTGTTAGCGCGAG	1199
Db	1141	GTGCGCTTCACCGAGCCCCCGCCAGCGTGTCTGAGCTCTGGCGCGCTGTGTTAGCGCGAG	1200
QY	1200	CACAGCTCGGAGTAAAGGAAAGGATGATGTGCGCC	1235
Db	1201	CACAGCTC3GAGTAAAGGAAAGGATGATGTGCGCC	1236

	RESULT	9			
	LOCUS	A87607			
	DEFINITION	A87607 Sequence 4 from Patent WO9836089.	PAT		22-JAN-2000
	ACCESSION	A87607			
	VERSION	A87607.1			
	KEYWORDS	GI:5736247			
	SOURCE	.			
	ORGANISM	unidentified.			
	REFERENCE	unidentified			
	AUTHORS	1 (bases 1 to 1237)			
	TITLE	Flohe,L and Singh,M.			
	JOURNAL	TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE			
		Patent: WO 9836089-A 20-AUG-1998;			
		FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE),			
		Location/Qualifiers			
FEATURES		1..1237			
source		/organism="unidentified"			
		/db_xref="taxon:32644"			
BASE COUNT		236 a 394 c 386 g 221 t			
ORIGIN					

Dp	361	CGTGTTCACACGATGGTGTGTTGGATTCCGGCACACGTCATATGGCTTACGAGACCGTC	420
Qy	420	CAGACCGCCGAAAGCGGCACTACCCCTGCTTGGCCCCGATGAGCGAAAGTCGGCGGTGCACATC	479
Dp	421	CGAGCGCCCGACGGCGCACTACCCCTGCTTGGCCCCGATGAGCGAAAGTCGGCGGTGCACATC	480
Qy	480	GCCGCCAGGTTGGCGCTTACCACTGATGGCAACCCAGAGGGGCGCGGTGTGCTGATG	539
Dp	481	GCCGCCAGGTTGGCGCTTACCACTGATGGCAACCCAGAGGGGCGCGGTGTGCTGATG	540
Qy	540	GGCGGGGGTGGCCGGCGTCGAAACCGCGCCGACGTGTGGTATTCGGCGCGCGGACACCGCGGC	599
Dp	541	GGCGGGGGTGGCCGGCGTCGAAACCGCGCCGACGTGTGGTATTCGGCGCGCGGACACCGCGGC	600
Qy	600	TACAACGCGACGCCCATTCGCCCAACGGCGATGGGCGCGACCGCTTACGGTTCTAGACATCAAC	659
Dp	601	TACAACGCGACGCCCATTCGCCCAACGGCGATGGGCGCGACCGTTACGGTTCTAGACATCAAC	660
Qy	660	ATCGACAAACTTCGGGCAATCTGACAGCCGAGTTCTGGGGCCGGATCCACACTCGCTACTCA	719
Dp	661	ATCGACAAACTTCGGGCAATCTGACAGCCGAGTTCTGGGGCCGGATCCACACTCGCTACTCA	720
Qy	720	TCGGGCTACGAGCTCGAGGGGGCGCTGAAACGTTGCCGACCTCGATGATTGGGCGCGCTG	779
Dp	721	TCGGGCTACGAGCTCGAGGGGGCGCTGAAACGTTGCCGACCTCGATGATTGGGCGCGCTG	780
Qy	780	GTGCGCAGGGCGCCAGGCGACCCCAATTACTCTCGAATTCACTGTGCGGCATATGAACA	839
Dp	781	GTGCGCAGGGCGCCAGGCGACCCCAATTACTCTCGAATTCACTGTGCGGCATATGAACA	840
Qy	840	GGTCGCGTACTGATGGATATAGCCATCGACGACGAGGGCGGCTTTCCAAAGGTCACAGCG	899
Dp	841	GGTCGCGTACTGATGGATATAGCCATCGACGACGAGGGCGGCTTTCCAAAGGTCACAGCG	900
Qy	900	ACCACTACGACACCCCGACGTTGCCGTGCACGACACGCTGTTTTACTGCTGGCGAAC	959
Dp	901	ACCACTACGACACCCCGACGTTGCCGTGCACGACACGCTGTTTTACTGCTGGCGAAC	960
Qy	960	ATGCGCCGCTGGGGGCGGAGAGAGCTGACCTACGGCGGTGACCAAGCGAGAGTGGCGTAT	1011
Dp	961	ATGCGCCGCTGGGGGCGGAGAGAGCTGACCTACGGCGGTGACCAAGCGAGAGTGGCGTAT	1020
Qy	1020	GTGCTCGAGCTTGCAGCACTATGGCTGGGGGGCGCGTGGCCGATCGAATCCGGCATTAGCC	1079
Dp	1021	GTGCTCGAGCTTGCAGCACTATGGCTGGGGGGCGCGTGGCCGATCGAATCCGGCATTAGCC	1080
Qy	1080	AAAGGCTTTTCGAGCGACGCAAGAGGGCGCTTACTTCGGAACGGGTGGCCACCGACTGGGG	1133
Dp	1081	AAAGGCTTTTCGAGCGACGCAAGAGGGCGCTTACTTCGGAACGGGTGGCCACCGACTGGGG	1144
Qy	1140	GTGCGCTTACCGAGAGCCCGCCAGCGTGTGGCTGACTTCGGCGCCCTGTTACGCGAG	1199
Dp	1141	GTGCGCTTACCGAGAGCCCGCCAGCGTGTGGCTGACTTCGGCGCCCTGTTACGCGAG	1200
Qy	1200	CACACGTGCGGAGCTAAGGGAAGGATGATTCGGCC	1235
Dp	1201	CACACGTGCGGAGCTAAGGGAAGGATGATTCGGCC	1236

	RESULT	10	
	AB9747		
LOCUS	AB9747	1237 bp	DNA
DEFINITION	Sequence 4 from Patent WO9832862.		PAT
ACCESSION	AB9747		
VERSION	AB9747.1	GI:6738281	
KEYWORDS	.		
SOURCE	unidentified.		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 1237)		
AUTHORS	Flohe,L. and Singh,N.		
TITLE	L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM		

JOURNAL Patent: WO 9832862-A 30-JUL-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES Location/Qualifiers
 Source 1..1237
 /orgnism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 236 a 394 c 386 g 221 t
 ORIGIN

Query Match 98.9%; Score 1220.8; DB 5; Length 1237;
 Best Local Similarity 99.8%; Pred. No. 2.9e-166;
 Matches 1233; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```

QY 1 ATCTTGAGATTAATGAACTTTCTTACACATGAAGGTACAGTATGAGAGGGTAAATC 60
DB 1 ATCTTGAGATTAATGAACTTTCTTACATGAGGTACAGTATGAGAGGGTAAATC 60
QY 61 ATGCGGTCGCTATTCGACGACGACCAAAAACAAAGAAATCCGGGTGGCATCACCCG 120
DB 61 ATGCGGTCGCTATTCGACGACGACCAAAAACAAAGAAATCCGGGTGGCATCACCCG 120
QY 121 GCCGGGTCGCGAAGTAAACCCGTCGTGCGCATGAGTGTCTCATCCAGGAGGTGCGGA 180
DB 121 GCCGGGTCGCGAAGTAAACCCGTCGTGCGCATGAGTGTCTCATCCAGGAGGTGCGGA 180
QY 181 GAGGCTCGCTATTCACGACGCGGATTTCAAGGGGCGAGCGCGCACTGTGCGCAC 240
DB 181 GAGGCTCGCTATTCACGACGCGGATTTCAAGGGGCGAGCGCGCACTGTGCGCAC 240
QY 241 GCCGACAGGTGAGGCGCGACGCTGATTTATGCTCAAGGTCAAAAGACCGATAGGGGG 300
DB 241 GCCGACAGGTGAGGCGCGACGCTGATTTATGCTCAAGGTCAAAAGACCGATAGGGGG 300
QY 301 GAATACGGCCCTCGGACACAGGCGG-GATCTTGTACGTTTCTTGTGATTTGGCGCGTCA 359
DB 301 GAATACGGCCCTCGGACACAGGCGG-GATCTTGTACGTTTCTTGTGATTTGGCGCGTCA 359
QY 360 CGTGTTCGACCATGCGTTTGTGATTTCCGGGACACAGTCAATTCCTTACACAGCCGTC 419
DB 360 CGTGTTCGACCATGCGTTTGTGATTTCCGGGACACAGTCAATTCCTTACACAGCCGTC 419
QY 420 CAGACGCGCAAGGCGACATACCCGTCGTCGCGGATGAGCAAGTCCGCGGTGACACT 479
DB 420 CAGACGCGCAAGGCGACATACCCGTCGTCGCGGATGAGCAAGTCCGCGGTGACACT 479
QY 480 GCCGCGCAGGTTGGCGCTTACACACTGATGCGAAAGGCGCGCGGTGTGTGATG 539
DB 480 GCCGCGCAGGTTGGCGCTTACACACTGATGCGAAAGGCGCGCGGTGTGTGATG 539
QY 540 GGGGGGTCGCGCGGCTGCAACCGCGCGGCGGCGTGTGATGCGCGCGCGCGCGGCG 599
DB 540 GGGGGGTCGCGCGGCTGCAACCGCGCGGCGGCGTGTGATGCGCGCGCGCGCGGCG 599
QY 600 TACAAGCGAAGCCGCGATCGCAACGCGGATGCGCGGCGGCTTACGCTTGTAGACATCAAC 659
DB 600 TACAAGCGAAGCCGCGATCGCAACGCGGATGCGCGGCGGCTTACGCTTGTAGACATCAAC 659
QY 660 ATGCAAACTTGGCACTCGACGCGGATTTGCGGCG GATCAGCACTGCTCTA 719
DB 660 ATGCAAACTTGGCACTCGACGCGGATTTGCGGCGGATTCACACTGCTCTA 719
QY 720 TCGGCTTACGAGTTCGAGGTCGCTCAACAGTGCAGACTGCTGTGAGTGGGCGGCTG 779
DB 720 TCGGCTTACGAGTTCGAGGTCGCTCAACAGTGCAGACTGCTGTGAGTGGGCGGCTG 779
QY 780 GTGCGCAGGCGCGACGCGCAATTAATGCTGGAATTCACCTGTGCGCATATGAACA 839
DB 780 GTGCGCAGGCGCGACGCGCAATTAATGCTGGAATTCACCTGTGCGCATATGAACA 839
QY 840 GGTGGGCTAGTGTGATGATGACGACGAGGCGGCTGTTGCAAGGCTCAGACGCG 899
DB 840 GGTGGGCTAGTGTGATGATGACGACGAGGCGGCTGTTGCAAGGCTCAGACGCG 899

```

```

QY 900 AACACCTACGACACCGGACGTTCCGCGCGACGACGACGCTTTACTGCTGGCGAAC 959
DB 901 AACACCTACGACACCGGACGTTCCGCGCGACGACGACGCTTTACTGCTGGCGAAC 960
QY 960 ATGCGCGCTCGGTGCGGACGAGTGCAGCTACGCGGCTGACCAAGCAGATGCGGTAT 1019
DB 961 ATGCGCGCTCGGTGCGGACGAGTGCAGCTACGCGGCTGACCAAGCAGATGCGGTAT 1020
QY 1020 GTGCTTGACCTTCCGACATGCTGCGGCGGCGGCTGCGGCTGCAATTCGCGACTAGCC 1079
DB 1021 GTGCTTGACCTTCCGACATGCTGCGGCGGCGGCTGCGGCTGCAATTCGCGACTAGCC 1080
QY 1080 AAAGGCTTTTCAGCAGCAGGAGGCGGCTGCTGCTGCGGAGGCGGCGGCGGCGGCGG 1139
DB 1081 AAAGGCTTTTCAGCAGCAGGAGGCGGCTGCTGCTGCGGAGGCGGCGGCGGCGGCGG 1140
QY 1140 GTGCGCTTACGAGCGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1199
DB 1141 GTGCGCTTACGAGCGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
QY 1200 CACAGCTCGGAGTAAGGAGGAGGAGGATGATGTCGGCC 1235
DB 1201 CACAGCTCGGAGTAAGGAGGAGGAGGATGATGTCGGCC 1236

```

RESULT 11

MTV002 56414 bp DNA BCT 17-JUN-1998
 LOCUS Mycobacterium tuberculosis H37Rv complete genome; segment 122/162.
 ACCESSION AL008967
 VERSION AL008967.1 GI:3261491
 KEYWORDS

SOURCE

ORGANISM Mycobacterium tuberculosis.
 Bacteria; Firmi; tes; Actinobacteriia; Ac. Inobacteriidae;
 Actinomyetales; Corynebacteriineae; Mycobacteriaceae;

REFERENCE

AUTHORS Cole, S.T., Brosch, R., Parkhill, J., Garier, T., Churcher, C., Harris, D., Gordon, S.V., Eigmeier, K., Gas, S., Barry III, C.E., Teke, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., O'Brien, A., Osborne, J., Quail, M.A., Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skelton, S., Squares, S., Squires, R., Sultson, J.E., Taylor, K., Whitehead, S. and Barrall, B.G.
 1 (bases 1 to 56414)

TITLE

Deciphering the biology of Mycobacterium tuberculosis from the

Complete genome sequence

Nature 393 (6685), 537-544 (1998)

MEDLINE

98295987

REMARK

Erratum: [[published erratum appears in Nature 1998 Nov

12:396(6707):190]]

2 (bases 1 to 56414)

ERRATUM

12:396(6707):190]]

ERRATUM

12:396(6707):190]]

ERRATUM

12:396(6707):190]]

ERRATUM

12:396(6707):190]]

ERRATUM

12:396(6707):190]]

ERRATUM

12:396(6707):190]]

CAUTION: In some cases we may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, or ttg) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

FEATURES

```

source
    1. 56414
    /organism="Mycobacterium tuberculosis"
    /strain="H37Rv"
    /db_xref="taxon:1773"
    <1..>56414
    /note="fragment designated v002. Does not represent a physical clone"
    <1..233
    /organism="Mycobacterium tuberculosis"
    /strain="H37Rv"
    /db_xref="taxon:1773"
    /clone="Y154"
    complement(3..527)
    /gene="recX"
    /gene="recX"
    complement(3..527)
    /note="Rv2738c, (MTV002.01c), len: 174 aa. recX, similar to eg. RECX PSEAE P37860 regulatory protein recX from Pseudomonas aeruginosa (153 aa), fasta scores: opt: 161 z-score: 257.2 E(): 3.6e-07, 30.7% identity in 137 aa overlap. Overlaps and extends CDS from overlapping cosmid MTC154.16c"
    /codon_start=1
    /transl_table=11
    /product="recX"
    /protein_id="CAI5532.1"
    /db_xref="GI:2624257"
    /db_xref="SPTREMBL:O33280"
    /translation="MTVSCPPSTSEEEQARALCLLTARSTRRAELAGOLAKRGY PEDGNVLDRLAAGLVYDIDFAEOMVQSRANAAKSRALAEHLAKVDDVITV LGGIDGAERGRAEKLVARARREVLIIDGTDEARVSRRLVAMLARVGGQLACEV VIELAEERRRRV"
    complement(493..2865)
    /gene="recA"
    complement(493..2865)
    /note="recA"
    /note="Rv2737c, (MTV002.02c), len: 790 aa. recA, identical to RECA MYCTU P26345 recA protein (790 aa). Contains self-splicing protein element (intron) from 2294 to 669 (c), similar to intron II from PR:E32317 (EMBL:Y13030) DNA-directed DNA polymerase (EC 2.7.7.7) from Thermococcus sp. (1829 aa), fasta scores: opt: 81 z-score: 235.2 E(): 6e-06, 24.6% identity in 183 aa overlap. Contains PS00017 ATP/GTP-binding site motif A (P-loop), PS00321 recA signature, and PS00881 protein splicing signature. See Davis et al, (1992) Cell 71(2):201-210"
    /codon_start=1
    /transl_table=11
    /product="recA"
    /protein_id="CAI5533.1"
    /db_xref="GI:2624258"
    /db_xref="SWISS-PROT:P26345"
    /translation="MTQTPDREKALAAVAOIKSYGKGSVMRLGDEARPIVITG SIALDVALIGTGLPGRVITIGPESGKTTVALHVAANAQAAGVAALFIDAHADP DYAKKIGVDTSILVSPDTEGEALAEADLINSGLDIVIDVSVALVRALEGEK GDSHVGLQARLKSQALRKMTGALNNGSTAIIFINLRDKIGVSGPETTGKAKF YASRVDRVREVELTKDGTAVNVRTRKVKVKNKCLAEGRIFEDPTGTHRIDVDVG RKPDIHVAAKDCGTLHARPVSWFDGTRDVIIGLRAGGAIWATPDHVLTFYGBRA AGLRKGRVAQPRRFDFGDSAPIPADHARLRLIGSGRSGWGKMPININOR ALIDDTRIATATGCAAHPOGRSIAIHRGCRNGVADICQAGYKGLAKMTKIPN WREPEPLADIYNNLIFGLFESDGMVSRBTGRLRGVYTTTSQALHQJHMLLRFV GSTVYRDPLOKRPSTVNGRRIDSKHOFVEVRISGMDVATFAESVPMGPPRAALIQ ALPEATQGRRRSQATYLAEMTDVAINTLDERGVTAQEAAMIGVAGSDPRGAKOV LGASRLRDRVQALADLDKFLHMLAEELRSYREVLPTRARPTFLEVEELHTL VAGGVVHNSPPEKQAEFDIYKGISRSGSLDMGVGGLIRKSGAMFTYEGEGLS OGKFNARNFLVENADVADIEKKIKKELKLGAVVTDPSNDGVLPAVDF"

```

```

misc_feature
    complement(790..2112)
    /gene="recA"
    /note="recA intron"
    complement(790..807)
    /gene="recA"
    /note="PS00881 protein splicing signature"
    complement(2197..2223)
    /gene="recA"
    /note="PS00321 recA signature"
    complement(2644..2667)
    /gene="recA"
    /note="PS00017 ATP/GTP-binding site motif A (P-loop)"
    complement(2873..2877)
    /note="possible RBS upstream of Rv2737c"
    complement(3247..3453)
    /gene="Rv2738c"
    complement(3247..3453)
    /gene="Rv2738c"
    /note="Rv2738c, (MTV002.03c), len: 68 aa. Unknown, N-terminus is highly similar to the N-terminus of the upstream ORF MTV002.0/c (78.4% identity in 37 aa overlap); also similar to A020958|SC4H_5 Streptomyces coelicolor cosmid 4H8 (64aa) opt: 185 z-score: 283.5 E(): 2.9e-08; 39.7% identity in 63 aa overlap"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein Rv2738c"
    /protein_id="CAI5534.1"
    /db_xref="GI:2624260"
    /db_xref="SPTREMBL:O33281"
    /translation="MLAGVRLTFEHRVALHFGAAYSSVLLDHYLTGDFGRSAQAQI EDGVEPRDYRALCADFDVPHDRW"
    complement(3464..4630)
    /gene="Rv2739c"
    complement(3464..4630)
    /gene="Rv2739c"
    /note="Rv2739c, (MTV002.04c), len: 388 aa. Probable trans. use, similar to eg. TR:051560 (EMBL:L28170) Pseudomonas aeruginosa rhamnosyl transferase (426 aa), fasta scores: opt: 178 z-score: 226.3 E(): 1.9e-05, 25.9% identity in 425 aa overlap. Equivalent to Mycobacterium leprae protein MLCB33.02c (392 aa), fasta scores: 9P1294723|MLCB33.2 Mycobacterium leprae cosmid B33 opt: 2112 z-score: 2364.5 E(): 0; 80.9% identity in 388 aa overlap"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein Rv2739c"
    /protein_id="CAI5535.1"
    /db_xref="GI:2624261"
    /db_xref="SPTREMBL:O33282"
    /translation="WRVAVVAGPDGSHSPALICQPRRAADPTLTFTGVEMEAR AAGIDAVELDGLAATRDLDAGARLHRRAAQAVALNVRRLALEVELVSVITACGG MAELIGIPWELNHPILYLPSCGLPIPGSGLAAGTIGRRLDRLATLGRSBRG LRORAARVREIIGLPARDGPLRLIATLPALEVPDPDPAVAVVAGPLHEPPTRYLA IPAGGPVAVVAPSTALGTAGTLEVALQSLTIGPCTVSGSRVYSRISGADLYPPH AVAGLSQAEILTRADLVITCGGHRGMVAKTLLAGVPVYVGGGQWETIARVROGS AVILRPLTADLVAANVEVLSPPREARRRAASVAGADPVRCBDALALAG"
    4674..5123
    /gene="Rv2740"
    4674..5123
    /gene="Rv2740"
    /note="Rv2740, (MTV002.05), len: 149 aa. Unknown, similar to M. leprae hypothetical protein TR:049650 (EMBL:294723) MLCB33.03 (82235_53-140) (178 aa), fasta scores: opt: 498 z-score: 538.9 E(): 7.3e-23, 51.6% identity in 161 aa overlap"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein Rv2740"
    /protein_id="CAI5536.1"
    /db_xref="GI:2624262"
    /db_xref="SPTREMBL:O33283"

```

/translation="MAELTETSPETPETTEAIRAVEAFUNALQNEDEPTDALGDL
 VYENVGSRIRGRRATILLRMOGRVFEKIRIADGAVALTERDALIIGPLRV
 QFWVGEVDDGRITLWRDYFDYDMEKGLRLVALVLPSTLAL".
 gene 5355..6932
 /gene="PE_PGRS"
 5355..6932
 /gene="PE_PGRS"
 /note="RV2741, (MTV002.06), 1..: 525 aa; Member of M.
 tuberculosis PE_PGRS subfamily, similar to many eg.

Query Match 98.9%; Score 1220.8; DB 1; Length 56414;
 Best Local Similarity 99.8%; Pred. No. 1.2e-166;
 Matches 1233; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATCTTGAGATTATATCGAACTTTCTTCACTAGAGCGTACAGTACAGAGGGTAATC 60
 Db 38201 ATCTTGAGATTATATCGAACTTTCTTCACTAGAGCGTACAGTACAGAGGGTAATC 38260
 QY 61 ATGCGCGTTCGATTTCGACGAGACCAAAAACAAGAAATTCGGGTGCCATCACCCCG 120
 Db 38261 ATGCGCGTTCGATTTCGACGAGACCAAAAACAAGAAATTCGGGTGCCATCACCCCG 38320
 QY 121 GCCGCGTTCGAGAACTAACCGGTGCGCATGAGTGTCTATCCAGGAGGTGCGGGA 180
 Db 38321 GCCGCGTTCGAGAACTAACCGGTGCGCATGAGTGTCTATCCAGGAGGTGCGGGA 38380
 QY 181 GAGGCGTTCGCTATACCGAGCGGATTTCAAGCGGCGCAACGTGTCGCGCAC 240
 Db 38381 GAGGCGTTCGCTATACCGAGCGGATTTCAAGCGGCGCAACGTGTCGCGCAC 38440
 QY 241 GCCGCGTTCGAGTTCGCGGCGAGCGTATTTATTCGCAAAAGAAACCGATAGCGCG 300
 Db 38441 GCCGCGTTCGAGTTCGCGGCGAGCGTATTTATTCGCAAAAGAAACCGATAGCGCG 38500
 QY 301 GAATACGCGCGCTTCGACGAGCGGCG -GATCTTGTTCACGTTCTTCATTTGGCGCGTCA 359
 Db 38501 GAATACGCGCGCTTCGACGAGCGGCGAGATCTTGTTCACGTTCTTCATTTGGCGCGTCA 38560
 QY 360 CGTGTTCACCGATGCTTGTGATTCGCGCACACGTCATTTGCTTACGAGACCGCTC 419
 Db 38561 CGTGTTCACCGATGCTTGTGATTCGCGCACACGTCATTTGCTTACGAGACCGCTC 38620
 QY 420 CAGACCGCGGAGCGGCGCTACCGCTGCTTGGCGGAGGAAATTCGCGGCTGCAATC 479
 Db 38621 CAGACCGCGGAGCGGCGCTACCGCTGCTTGGCGGAGGAAATTCGCGGCTGCAATC 38680
 QY 480 GCCGCGCGAGTTCGCTTACCACTGATGCGAAACCAAGGCGCGCGGTGTGCTGATG 539
 Db 38681 GCCGCGCGAGTTCGCTTACCACTGATGCGAAACCAAGGCGCGCGGTGTGCTGATG 38740
 QY 540 GCGGCGGTCGCGGCGTTCGAGACCGGCGGAGTTCGCTGATGCGCGCGCGCGCG 599
 Db 38741 GCGGCGGTCGCGGCGTTCGAGACCGGCGGAGTTCGCTGATGCGCGCGCGCGCG 38800
 QY 600 TACAACGAGCGCGGCGATGCGCAAGCGGCGGAGCGGCTTACGCTTACGATCAAC 659
 Db 38801 TACAACGAGCGCGGCGATGCGCAAGCGGCGGAGCGGCTTACGCTTACGATCAAC 38860
 QY 660 ATCGCAAACTTCGCGCACTGAGCGCGGAGTTCGCG :CGATTCACACTGCGCTACT A 719
 Db 38861 ATCGCAAACTTCGCGCACTGAGCGCGGAGTTCGCGCGGATTCACACTGCGCTACTCA 38920
 QY 720 TCGGCTACGAGCTTCGAGGTCGCTAAAGTTCGCGGAGTTCGAGGTCGCTGCTG 779
 Db 38921 TCGGCTACGAGCTTCGAGGTCGCTAAAGTTCGCGGAGTTCGAGGTCGCTGCTGCTG 38980
 QY 780 GTGCGAGCGCGCAAGCGCAACCAATTAAGTTCGATTTGCTGCGCATATGAACCA 839
 Db 38981 GTGCGAGCGCGCAAGCGCAACCAATTAAGTTCGATTTGCTGCGCATATGAACCA 39040
 QY 840 GGTGCGTACTGAGTATAGCATGACGAGGCGGCTTTTCGAGGTCGCGACCG 899
 Db 39041 GGTGCGTACTGAGTATAGCATGACGAGGCGGCTTTTCGAGGTCGCGACCG 39100

QY 900 ACCACTACGACCCAGCTTGGCGTGCACGACACGCTGTTTACTGCGTGGCAAC 959
 Db 39101 ACCACTACGACCCAGCTTGGCGTGCACGACACGCTGTTTACTGCGTGGCAAC 39160
 QY 960 ATGCGCGCTTCGTCGCGGAGAGAGCTGACCTACGCGGTGACCAAGCGAGATGCGGTAT 1019
 Db 39161 ATGCGCGCTTCGTCGCGGAGAGAGCTGACCTACGCGGTGACCAAGCGAGATGCGGTAT 39220
 QY 1020 GTGCTGAGCTTCGCGACCATGCTGCGGCGGCGGTGCGGTGCAATTCGCGACTAGCC 1079
 Db 39221 GTGCTGAGCTTCGCGACCATGCTGCGGCGGCGGTGCGGTGCAATTCGCGACTAGCC 39280
 QY 1080 AAGGCTTCGAGCGACGAGAGGCGCTTACTGTCGCAAGGCGGTGCGCAAGGAGCTGGGG 1139
 Db 39281 AAGGCTTCGAGCGACGAGAGGCGCTTACTGTCGCAAGGCGGTGCGCAAGGAGCTGGGG 39340
 QY 1140 GTGCGCTTCACGAGCGCGCGAGCTGCTGCGCTGACTTCGCGCGCTGCTACGCGAG 1199
 Db 39341 GTGCGCTTCACGAGCGCGCGAGCTGCTGCGCTGACTTCGCGCGCTGCTACGCGAG 39400
 QY 1200 CACACGTTCGAGGAGTAAAGGAGGAGGATGTCGCGC 1235
 Db 39401 CACACGTTCGAGGAGTAAAGGAGGAGGATGTCGCGC 39436

RESULT 12

AB7606 1235 bp DNA PAT 22-JAN-2000
 LOCUS
 DEFINITION Sequence 3 from Patent WO9836089.
 ACCESSION AB7606
 VERSION AB7606.1 GI:5736246
 KEYWORDS

SOURCE
 ORGANISM
 unclassified.
 unclassified.

REFERENCE
 1 (bases 1 to 1235)
 Flohe/L. and Singh,M.

TITLE
 TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
 JOURNAL
 Patient: WO 9836089-A 20-AUG-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)

FEATURES
 source
 1..1235
 location/Qualifiers

BASE COUNT 236 a 394 c 382 g 220 t 3 others
 ORIGIN

Query Match 98.7%; Score 1218.4; DB 5; Length 1235;
 Best Local Similarity 99.5%; Pred. No. 6.4e-166;
 Matches 1230; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1 ATCTTGAGATTATATCGAACTTTCTTCACTAGAGCGTACAGTACAGAGGGTAATC 60
 Db 1 ATCTTGAGATTATATCGAACTTTCTTCACTAGAGCGTACAGTACAGAGGGTAATC 60
 QY 61 ATGCGCGTTCGATTTCGACGAGACCAAAAACAAGAAATTCGGGTGCCATCACCCCG 120
 Db 61 ATGCGCGTTCGATTTCGACGAGACCAAAAACAAGAAATTCGGGTGCCATCACCCCG 120
 QY 121 GCCGCGTTCGAGAACTAACCGGTGCGCATGAGTGTCTATCCAGGAGGTGCGGGA 180
 Db 121 GCCGCGTTCGAGAACTAACCGGTGCGCATGAGTGTCTATCCAGGAGGTGCGGGA 180
 QY 181 GAGGCGTTCGCTATACCGAGCGGAGATTTCAAGGCGGAGCGCGCAACTGTCGCGCAC 240
 Db 181 GAGGCGTTCGCTATACCGAGCGGAGATTTCAAGGCGGAGCGCGCAACTGTCGCGCAC 240
 QY 241 GCGGACGAGTTCGCGGCGGCGCTGATTTTCTCAAGTCAAGGATGCGGCGG 300
 Db 241 GCGGACGAGTTCGCGGCGGCGCTGATTTTCTCAAGTCAAGGATGCGGCGG 300

Qy	301	GAAGAGCGCGCGCTGGAGACAGCGGGC-GA	CTGTTACAGCTTTGTCATTTGGCCGCGCA	359
Db	301	GAATACGGCGCGCTGGAGACAGGGGAGATCT	TTGTTACGCTTTTCATTTGGCCGCGTCA	360
Qy	360	CGTCTCTTGCACCGATGGCTGTTGGATTCCG	GACACAGCTCAATTTGCTACAGACGTC	419
Db	361	CGTCTTGCACCGATGGCTGTTGGATTCCGGAC	CCACGATCTCAATTTGCTACAGACGTC	420
Qy	420	CAGACCGCCGAAAGGCGCATACCCCTGCTTGC	CCCGATAGGAAGTGGCCGGTGCAGTC	479
Db	421	CAGACCGCCGAGCGCATACCCCTGCTTGCCTT	CCGATAGGAAGTGGCCGGTGCAGTC	480
Qy	480	GCCGCCAGGTTGGCGCTTACCACTGATGCGAA	CCCAAGGGGGCGCGGTGTGCTATG	539
Db	481	GCCGCCAGGTTGGCGCTTACCACTGATGCGAA	CCCAAGGGGGCGCGGTGTGCTATG	540
Qy	540	GCGCGGGTGGCGCGCTGCAGACCGCGCGACG	TGCTGTGATGGCGCGCGCGACCGCGCGC	599
Db	541	GCGCGGGTGGCGCGCGCTGCAGACCGCGCGAC	TGCTGTGATGGCGCGCGCGACCGCGCGC	600
Qy	600	TACACGCGAGCGCCGCTATCGCCAAAGCGCAT	GGCGCGCACGCTTACGCTTATAGACATCAAC	659
Db	601	TACACGCGAGCGCCGCTATCGCCAAAGCGCAT	GGCGCGCACGCTTACGCTTATAGACATCAAC	660
Qy	660	ATCGACAAACTTGGCGCACTGCAGCGCGAGTT	CTGGCGCGCGATCCACACTGCTACTCA	719
Db	661	ATCGACAAACTTGGCGCAACTGCAGCGCGAGTT	CTGGCGCGCGATCCACACTGCTACTCA	720
Qy	720	TCGGCTCTACGAGCTCGAGGGGCGCGTCAAA	CGTGCAGACTGGATTTGGGGCGCTCTGTG	779
Db	721	TCGGCTCTACGAGCTCGAGGGGCGCGTCAAA	CGTGCAGACTGGATTTGGGGCGCTCTGTG	780
Qy	780	GTCGCCAGCGCCCAAGGACCCCAATTAGTCTG	GAATTCACTGTGTGCGCATATGAACCA	839
Db	781	GTCGCCAGCGCCCAAGGACCCCAATTAGTCTG	GAATTCACTGTGTGCGCATATGAACCA	840
Qy	840	GTCGCGTACTGATGATATTAAGCATGACACG	GGGGGGGCTTTGGAAGGCTCAGACGCG	899
Db	841	GTCGCGTACTGATGATATTAAGCATGACACG	GGGGGGGCTTTGGAAGGCTCAGACGCG	900
Qy	900	ACCACTTACGACCAACCCGACGTTTGCCCTG	CACAGACGCTGTTTTACTGCGTGGCGAAC	959
Db	901	ACCACTTACGACCAACCCGACGTTTGCCCTG	CACAGACGCTGTTTTACTGCGTGGCGAAC	960
Qy	960	ATGCGCGCCTCGTGTGCCGAGAACGCTGCAC	TACGCTACGCAAGCGGACGATGGCGAT	1019
Db	961	ATGCGCGCCTCGTGTGCCGAGAACGCTGCAC	TACGCTACGCAAGCGGACGATGGCGAT	1020
Qy	1020	GTCGCTCGAGCTTGGCGACATGCTGTGCGGG	CGCGCTGCGAGTCCGGCACTACACC	1079
Db	1021	GTCGCTCGAGCTTGGCGACATGCTGTGCGGG	CGCGCTGCGAGTCCGGCACTACACC	1080
Qy	1080	AAAGGCTCTTTGACGCGACGAAGGGGGCTTAC	TGTCGGAAGGGTGGCCACGCACTGGGG	1139
Db	1081	AAAGGCTCTTTGACGCGACGAAGGGGGCTTAC	TGTCGGAAGGGTGGCCACGCACTGGGG	1140
Qy	1140	GTCGCGCTTACAGAGCGCGCGGAGGTGCTG	AGCTCGGCGCTGCTTAAAGCCGAG	1199
Db	1141	GTCGCGCTTACAGAGCGCGCGGAGGTGCTG	AGCTCGGCGCTGCTTAAAGCCGAG	1200
Qy	1200	CACACGTGCGGAGTAAAGGAAGCATGATGTG	CGGC 1234	
Db	1201	CACACNTGCGGAGTAAAGGAAGCATGATGTG	CNC 1235	
RESULT 13				
LOCUS	AB9746	1235 bp	DNA	PAT
DEFINITION	Sequence 3 from Patent WO9832862.			
ACCESSION	AB9746			22-JAN-2000
VERSION	AB9746.1	GI:6738280		
KEYWORDS	unidentified.			
SOURCE				

ORGANISM	unidentified. unclassified.	1 (bases 1 to 1235)
REFERENCE	Flohe, L. and Singh, M.	
AUTHORS	L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM	
TITLE	Patent: WO 9832862-A 30-JUL-1998;	
JOURNAL	FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)	
FEATURES	Location/Qualifiers	
source	1. 1235	
BASE COUNT	236 a 394 c 382 g 220 t 3 others	
ORIGIN		
Query Match	98.7% Score 1218.4; DB 5; Length 1235;	
Best Local Similarity	99.6% Pred. No. 6.4e-166;	
Matches 1230; Conservative	0: Mismatches 4; Indels 1; Gaps 1	
QY	1 ATCTTGAGATTAAATGAACTTCTTACACACTGAAGCGCTACAGTATGAGAGGGGTATC	60
Db	1 ATCTTGAGATTAAATGAACTTCTTACACACTGAAGCGCTACAGTATGAGAGGGGTATC	60
QY	61 ATGGCGCTCGGTATTCGCCAGCCGAGACCAAAACAGAAATTCGGGGGCCATCACCCG	120
Db	61 ATGGCGCTCGGTATTCGCCAGCCGAGACCAAAACAGAAATTCGGGGGCCATCACCCG	120
QY	121 GCCGCGCTCGCGGAACTAACCCGCTGTGGCCATGAGTGCTATCATCAGGCGAGTGGCGGA	180
Db	121 GCCGCGCTCGCGGAACTAACCCGCTGTGGCCATGAGTGCTATCATCAGGCGAGTGGCGGA	180
QY	181 GAGGGCTCGGCTATCACCGACGGGATTTCAAGGGCGGACGGCGCAACTGTGGGAC	240
Db	181 GAGGGCTCGGCTATCACCGACGGGATTTCAAGGGCGGACGGCGCAACTGTGGGAC	240
QY	241 GCCGACGAGGTGAGGGCGGCGCGCTGATTTATGTCACAGGTCAAAAGAACGATAGCGGG	300
Db	241 GCCGACGAGGTGAGGGCGGCGCGCTGATTTATGTCACAGGTCAAAAGAACGATAGCGGG	300
QY	301 GAATAGCGCGCGCTGCGACACGGGG- GATCTTGTTCAGTCTTTCATTTGGCCGGCTCA	359
Db	301 GAATAGCGCGCGCTGCGACACGGGG- GATCTTGTTCAGTCTTTCATTTGGCCGGCTCA	359
QY	360 CGTGCTTGACCCATGCTGTGTGATTCGGGACACAGCTCAATTCCTACGAAACCGTC	419
Db	360 CGTGCTTGACCCATGCTGTGTGATTCGGGACACAGCTCAATTCCTACGAAACCGTC	419
QY	361 CGTGCTTGACCCATGCTGTGTGATTCGGGACACAGCTCAATTCCTACGAAACCGTC	420
Db	361 CGTGCTTGACCCATGCTGTGTGATTCGGGACACAGCTCAATTCCTACGAAACCGTC	420
QY	420 CAGACCGCGGAGCGGACACTCCCTGCTTGGCCCGGATGAGGGAAGTGGCCGTGCACTC	479
Db	421 CAGACCGCGGAGCGGACACTCCCTGCTTGGCCCGGATGAGGGAAGTGGCCGTGCACTC	480
QY	480 GCGGCCCAAGTGTGGCGTTACCACTGATGCGAACCAAGGGGGCCGGGTGTCTATG	539
Db	481 GCGGCCCAAGTGTGGCGTTACCACTGATGCGAACCAAGGGGGCCGGGTGTCTATG	540
QY	540 GCGGGGGTGGCCCGCGTCGAACCGGCGGACGTCGTGTATGAGGGCGGACCGCGGG	599
Db	541 GCGGGGGTGGCCCGCGTCGAACCGGCGGACGTCGTGTATGAGGGCGGACCGCGGG	600
QY	600 TACAACGCGACCGCATGCGCAACGGGATGGCGCGACCGTTACGGTCTTAGACATCAAC	659
Db	601 TACAACGCGACCGCATGCGCAACGGGATGGCGCGACCGTTACGGTCTTAGACATCAAC	660
QY	660 ATTCGACAAACTTGGGCACTCGACGCCGAGTTCTGGGGCCGATCCACACTCGCTACTCA	719
Db	661 ATTCGACAAACTTGGGCACTCGACGCCGAGTTCTGGGGCCGATCCACACTCGCTACTCA	720
QY	720 TCGGCTCTAGAGCTCGAGGGTGGCGTCAAAACGTCGCGACTGGTATTTGGGGCGGTCTG	779
Db	721 TCGGCTCTAGAGCTCGAGGGTGGCGTCAAAACGTCGCGACTGGTATTTGGGGCGGTCTG	780
QY	780 GTGCAGGCGCCAAAGCACCACAAATTAGTCTGGAATTCACTTGTGCGCATATGAAACCA	839

```

Db 781 GNGCAGGCGCCAGGACCCCAATTAGTCTGCAATTCATTGTGCGCGCATATGAACCA 840
QY 840 GGTGGGTAATCTGTGATATACCATCGACGAGGGGGTGTGTTGAAAGCTCAGACCG 899
Db 841 GGTGGGTAATCTGTGATATACCATCGACGAGGGGGTGTGTTGAAAGCTCAGACCG 900
QY 900 ACCACTACGACACCGCGTTCGGCGGCGACGACGCTGTTTACTGCGTGGGGAAC 959
Db 901 ACCACTACGACACCGCGTTCGGCGGCGACGACGCTGTTTACTGCGTGGGGAAC 960
QY 960 ATGCCCGCTCGGTGCGGAAGAGCTGACCTACGGCTGACCAAGCGACGATGCCGTAT 1019
Db 961 ATGCCCGCTCGGTGCGGAAGAGCTGACCTACGGCTGACCAAGCGACGATGCCGTAT 1020
QY 1020 GTGCTCGACCTTCCGACCATGAGTGGGCGGGGCTGCGGCTCGAATCCGGCAGTACGC 1079
Db 1021 GTGCTCGACCTTCCGACCATGAGTGGGCGGGGCTGCGGCTCGAATCCGGCAGTACGC 1080
QY 1080 AAGAGCTCTTCGACGACGAAAGGGGCGTACTGTCCAAAGGATGCGACGACCTGGGG 1139
Db 1081 AAGAGCTCTTCGACGACGAAAGGGGCGTACTGTCCAAAGGATGCGACGACCTGGGG 1140
QY 1140 GTGCGCTTACCGAGCGCCGACAGCTGTGCGCTGACTCTCGCGCGCTGTTACGCCGAG 1199
Db 1141 GTGCGCTTACCGAGCGCCGACAGCTGTGCGCTGACTCTCGCGCGCTGTTACGCCGAG 1200
QY 1200 CACACTCGGAGTAAGGAGGATGATGCGGC 1234
Db 1201 CACACTCGGAGTAAGGAGGATGATGCGNC 1235

```

RESULT 14

```

LOCUS A87608 1228 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 5 from Patent WO936089.
ACCESSION A87608
VERSION A87608.1 GI:6736248
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1228)
AUTHORS Flohe, L. and Singh, M.
TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL Patent: WO 936089-A-20-AUG-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
source 1..1228
/organism="unidentified"
/db_xref="taxon:32644"

```

```

BASE COUNT 236 a 391 c 382 g 219 t
ORIGIN

```

Query Match

Best Local Similarity 98.2%; Score 1212.8; DB 5; Length 1228;

Matches 1225; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```

QY 1 ATCTTCAGATTAATCAACTTCTTACACTGAAACCTACAGTATGAGAGGGTAAATC 60
Db 1 ATCTTCAGATTAATCAACTTCTTACACTGAAACCTACAGTATGAGAGGGTAAATC 60
QY 61 ATGCGGTGGGTATTCGACCGGAGCAAAACAAAGATTCGGGGTGGCATCACCGCG 120
Db 61 ATGCGGTGGGTATTCGACCGGAGCAAAACAAAGATTCGGGGTGGCATCACCGCG 120
QY 121 GCGGCGTGGCGAACTAACCCGCTGTGGCCATGAGGTGCTATCCAGGCAAGTGGCGGA 180
Db 121 GCGGCGTGGCGAACTAACCCGCTGTGGCCATGAGGTGCTATCCAGGCAAGTGGCGGA 180
QY 181 GAGGCTGGGCTATCAACCGAGCGGATTTCAAGGCGGC/GCGCGCAACTGTGGCGAC 240
Db 181 GAGGCTGGGCTATCAACCGAGCGGATTTCAAGGCGGC/GCGCGCAACTGTGGCGAC 240

```

```

QY 241 GCCGACAGCTGTGGCGCGAGCGTATTATTGCTCAAGTCGAAGAACGATAGCGCG 300
Db 241 GCCGACAGCTGTGGCGCGAGCGTATTATTGCTCAAGTCGAAGAACGATAGCGCG 300
QY 301 GAATACGCGCGCTCGACACAGCGGC-GATCTGTTCAGCTTCGATTTGGCGCGTCA 359
Db 301 GAATACGCGCGCTCGACACAGCGGCAGATCTTGTTCAGCTTCGATTTGGCGCGTCA 360
QY 360 CGTCTTCACACCGATGCGTGTGGATTCGCGACACGATCAATTGCTACGAGACGTC 419
Db 361 CGTCTTCACACCGATGCGTGTGGATTCGCGACACGATCAATTGCTACGAGACGTC 420
QY 420 CAGACGCGCGAAGGGCGCATACCCCTGCTGGCCCGATGAGGAGTGGCGGCTGATC 479
Db 421 CAGACGCGCGAAGGGCGCATACCCCTGCTGGCCCGATGAGGAGTGGCGGCTGATC 480
QY 480 GCGGCGCGAGTTGGCGCTTACACCTGATGCGAACCCAAAGGGGGCGGGTGTGATG 539
Db 481 GCGGCGCGAGTTGGCGCTTACACCTGATGCGAACCCAAAGGGGGCGGGTGTGATG 540
QY 540 GCGGCGGTGCGCGCGTGCACACCGCGACGCTGTGTGATGCGCGCGACCGCGCG 599
Db 541 GCGGCGGTGCGCGCGCGTGCACACCGCGACGCTGTGTGATGCGCGCGACCGCGCG 600
QY 600 TCAACGCGAGCGCGCATGCGCAAGGGCATGGGCGGACGCTTACGAGTTCATGAC 659
Db 601 TCAACGCGAGCGCGCATGCGCAAGGGCATGGGCGGACGCTTACGAGTTCATGAC 660
QY 660 ATCGACAAACTTCGCGCACTGACGCGCGAGTCTGCGCGCGGATCGACACTGCTACTCA 719
Db 661 ATCGACAAACTTCGCGCACTGACGCGCGAGTCTGCGCGCGGATCGACACTGCTACTCA 720
QY 720 TCGGCTTACGAGCTGAGGGTGCCTCAAAAGTGCAGACCTGTGATTTGGGCGCTCTG 779
Db 721 TCGGCTTACGAGCTGAGGGTGCCTCAAAAGTGCAGACCTGTGATTTGGGCGCTCTG 780
QY 780 GTGCGAGGCGCGCAAGGCAATTAATGCTGCAATTAATGCTGCGCGATGAGAAACA 839
Db 781 GTGCGAGGCGCGCAAGGCAATTAATGCTGCAATTAATGCTGCGCGATGAGAAACA 840
QY 840 GGTGGGTAATCTGTGATATACCATCGACGAGGGGGTGTGTTGAAAGCTCAGACCG 899
Db 841 GGTGGGTAATCTGTGATATACCATCGACGAGGGGGTGTGTTGAAAGCTCAGACCG 900
QY 900 ACCACTACGACACCGCGCTTCGCGCTGACGACAGCGCTGTTTACTGCGTGGCGAAC 959
Db 901 ACCACTACGACACCGCGCTTCGCGCTGACGACAGCGCTGTTTACTGCGTGGCGAAC 960
QY 960 ATGCCCGCTCGGTGCGGAAGAGCTGACCTACGGCTGACCAAGCGACGATGCCGTAT 1019
Db 961 ATGCCCGCTCGGTGCGGAAGAGCTGACCTACGGCTGACCAAGCGACGATGCCGTAT 1020
QY 1020 GTGCTCGACCTTCCGACCATGAGTGGGCGGGGCTGCGGCTCGAATCCGGCAGTACGC 1079
Db 1021 GTGCTCGACCTTCCGACCATGAGTGGGCGGGGCTGCGGCTCGAATCCGGCAGTACGC 1080
QY 1080 AAGAGCTCTTCGACGACGAAAGGGGCGTACTGTCCAAAGGATGCGACGACCTGGGG 1139
Db 1081 AAGAGCTCTTCGACGACGAAAGGGGCGTACTGTCCAAAGGATGCGACGACCTGGGG 1140
QY 1140 GTGCGCTTACCGAGCGCCGACAGCTGTGCGCTGACTCTCGCGCGCTGTTACGCCGAG 1199
Db 1141 GTGCGCTTACCGAGCGCCGACAGCTGTGCGCTGACTCTCGCGCGCTGTTACGCCGAG 1200
QY 1200 CACACTCGGAGTAAGGAGGATGATGCGGC 1234
Db 1201 CACACTCGGAGTAAGGAGGATGATGCGGC 1235

```

RESULT 15

```

LOCUS A89748 1228 bp DNA PAT 22-JAN-2000

```

DEFINITION Sequence 5 from Patent WO9832862.
 ACCESSION A89748
 VERSION A89748.1 GI:6738282
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1228)
 AUTHORS Flohe L. and Singh M.
 TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
 JOURNAL Patent: WO 9832862-A 30-JUL-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source 1..1228
 Location/Qualifiers
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 236 a 391 c 382 g 219 t
 ORIGIN

Query Match 98.2% Score 1212.8; DB 5; Length 1228;
 Best Local Similarity 99.8%; Pred. No. 4e-165; 2; Indels 1; Gaps 1;
 Matches 1225; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATCTTGAGATTATGCAACTTTCTTCACACTGAGGCTACGATGAGAGGGTATC 60
 DB 1 ATCTTGAGATTATGCAACTTTCTTCACACTGAGGCTACGATGAGAGGGTATC 60
 QY 61 ATCGGGGTGGTATTCGAGCGAGACCAAAACAGCAATTCGGGGTGCATACCCCG 120
 DB 61 ATCGGGGTGGTATTCGAGCGAGACCAAAACAGCAATTCGGGGTGCATACCCCG 120
 QY 121 GCGGGCTCGGAGACTAACCGCTGCTGAGGAGTGCTCATCCAGGAGTGGCGGA 180
 DB 121 GCGGGCTCGGAGACTAACCGCTGCTGAGGAGTGCTCATCCAGGAGTGGCGGA 180
 QY 181 GAGGGCTCGGCTATCACCGACGGGATTTCAAGCGGCGGCGCACTGTGGCACC 240
 DB 181 GAGGGCTCGGCTATCACCGACGGGATTTCAAGCGGCGGCGCACTGTGGCACC 240
 QY 241 GCGGACGAGTGTGGGCGGAGCTGATTTATGCTCAAGGCTCAAGAAGCGATAGGGCG 300
 DB 241 GCGGACGAGTGTGGGCGGAGCTGATTTATGCTCAAGGCTCAAGAAGCGATAGGGCG 300
 QY 301 GAATAGGCGCGCTGGGACACGGGGC-GATCTTGTTCAGTTCTTGTGCGCGCTCA 359
 DB 301 GAATAGGCGCGCTGGGACACGGGGC-GATCTTGTTCAGTTCTTGTGCGCGCTCA 359
 QY 360 CGTGTTCGACCATGCTGTGTGATTCGGGACCACTGATTTGCTTACGAGACGCTC 419
 DB 361 CGTGTTCGACCATGCTGTGTGATTCGGGACCACTGATTTGCTTACGAGACGCTC 419
 QY 420 CAGACGCGGAGGCGGAGTACCTGCTGCGCGGATGAGCGAAGTGGCGGCTGACTC 479
 DB 421 CAGACGCGGAGGCGGAGTACCTGCTGCGCGGATGAGCGAAGTGGCGGCTGACTC 479
 QY 480 GCGGCGGAGTGTGGGCTTACCACTGATGCAACCAAGGGGGCGCGGTGTGCTGATG 539
 DB 481 GCGGCGGAGTGTGGGCTTACCACTGATGCAACCAAGGGGGCGCGGTGTGCTGATG 539
 QY 540 GGGGGGTGGCGGCGGCTGCAACCGGCGGAGCTGTGTGATTCGGGCGGCGGCGGCG 599
 DB 541 GGGGGGTGGCGGCGGCTGCAACCGGCGGAGCTGTGTGATTCGGGCGGCGGCGGCG 599
 QY 600 TACACGACGCGGAGTACGCAAGGAGTGGGGCGGCTTACGCTTACGATCAAC 659
 DB 601 TACACGACGCGGAGTACGCAAGGAGTGGGGCGGCTTACGCTTACGATCAAC 659
 QY 660 ATGACAAACTTGGGCACTGACGCGGAGTTCGGGCGGATCCCACTCGCTACTCA 719
 DB 661 ATGACAAACTTGGGCACTGACGCGGAGTTCGGGCGGATCCCACTCGCTACTCA 719
 QY 720 TCGGCTAGAGCTCGAGGCTGCGCTCAAAAGTGGCGACCTGGTATGGGGCGGCTCTG 779

DB 721 TCGGCTAGAGCTCGAGGCTGCGCTCAAAAGTGGCGACCTGGTATGGGGCGGCTCTG 780
 QY 780 GTGCGAGGCGGAGGCGGAGGCGGAGTGTGCTGAAATTCGCTGCGGCTATGAAGCA 839
 DB 781 GTGCGAGGCGGAGGCGGAGGCGGAGTGTGCTGAAATTCGCTGCGGCTATGAAGCA 840
 QY 840 GTGCGGCTAGTGTGATTTACCATGACGAGGCGGCTGTTTGAAAGGCTCAGACCG 899
 DB 841 GTGCGGCTAGTGTGATTTACCATGACGAGGCGGCTGTTTGAAAGGCTCAGACCG 900
 QY 900 ACCACCTAGACACCGGAGGCTGCGGCGGAGGAGGAGGCTGTTTACGCTGCGGAG 959
 DB 901 ACCACCTAGACACCGGAGGCTGCGGCGGAGGAGGAGGCTGTTTACGCTGCGGAG 960
 QY 960 ATGCGCGCTCGGTGCGGAGGAGGCTGAGGCTGAGGCTGAGGAGGAGGAGGCTAT 1019
 DB 961 ATGCGCGCTCGGTGCGGAGGAGGCTGAGGCTGAGGCTGAGGAGGAGGAGGCTAT 1020
 QY 1020 GTGCTGAGCTTGGCGGAGGAGGCTGAGGCTGAGGCTGAGGAGGAGGAGGCTAT 1079
 DB 1021 GTGCTGAGCTTGGCGGAGGAGGCTGAGGCTGAGGCTGAGGAGGAGGAGGCTAT 1080
 QY 1080 AAAGGTCTTTCGACGAGGAGGAGGCTGAGGCTGAGGAGGAGGAGGAGGAGGCTAT 1139
 DB 1081 AAAGGTCTTTCGACGAGGAGGAGGCTGAGGCTGAGGAGGAGGAGGAGGAGGCTAT 1140
 QY 1140 GTGCGTTCAGGAGGCGGCGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTAT 1199
 DB 1141 GTGCGTTCAGGAGGCGGCGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTAT 1200
 QY 1200 CACAGCTCGGAGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1227
 DB 1201 CACAGCTCGGAGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1228

Search completed: June 22, 2000, 14:56:13
 Job time: 17788 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 15:10:02 ; Search time 458.59 Seconds
(without alignments)
673.777 Million cell updates/sec

Title: US-09-362-485-8

Perfect score: 1235
Sequence: 1 ATCTTGCAGATTATCGAAC.....GGGAAGCGATGATCGGCC 1235

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Database: N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1206.4	97.7	1245	V49626	Mycobacterium tube
2	1206.4	97.7	1260	V49510	Mycobacterium sp.
3	1206.4	97.7	1260	V49625	Mycobacterium tube
4	454	36.8	682	V49511	Mycobacterium marl
5	265.2	21.5	1125	N91423	Sequence of heat-r
6	203.6	16.5	28171	V52155	Streptococcus pneu
7	166.4	13.5	9280	V74442	Staphylococcus aur
8	89.2	7.2	1074	T67971	H. pylori membrane
9	69.2	5.6	544	V43039	Streptococcus pneu
10	56.8	4.6	31096	V74370	Staphylococcus aur
11	49.2	4.0	390	Q21833	Randomising oligon
12	49.2	4.0	390	Q36859	PCR primer for 5'
13	48.8	4.0	15872	T68715	Streptomyces venez
14	46.4	3.8	985	V44439	Mycobacterium tube
15	46.4	3.8	985	V64548	M. tuberculosis im
16	45.4	3.7	833	O64203	snab gene encoding
17	45.4	3.7	5392	Q64201	Sequence compri
18	45	3.6	3946	T93610	Mycobacterium tube
19	44.8	3.6	535	V44537	M. tuberculosis im
20	44.8	3.6	535	V64537	M. tuberculosis im
21	44.4	3.6	17955	V56642	Actinoplanes sp. a
22	44.4	3.6	17955	T06769	Sorangium cellulos
23	43.6	3.5	28958	T89956	Sorangium cellulos
24	43.6	3.5	28958	T89956	The scraphen biosy
25	43.6	3.5	49377	V05287	Human adenosine A1
26	42.8	3.5	114955	X53491	smbR gene encoding
27	42.4	3.4	1833	O64206	Clone p5-4 encodin
28	42.4	3.4	2185	T08965	S. lividans protea
29	42.4	3.4	2186	T08965	Streptomyces freno
30	42.4	3.4	24379	T93095	Streptomyces roseo
31	42.4	3.4	24379	T93095	Mycobacterium tube
32	42.2	3.4	329	V44425	M. tuberculosis im
33	42.2	3.4	329	V64534	M. tuberculosis im
34	41.8	3.4	882	V44403	Mycobacterium tube

35	41.8	3.4	882	1	V64512	M. tuberculosis im
36	41	3.3	3946	1	T93610	Mycobacterium tube
37	41	3.3	29879	1	Q46806	eyA region of S.
38	40.8	3.3	2151	1	Q76252	HSV-2 protease/ICP
39	40.8	3.3	2151	1	Q76251	HSV-2 protease/ICP
40	40.8	3.3	2472	1	O84671	HSV-2 UL26 gene. N
41	40.8	3.3	20387	1	V62159	HSV-2 strain SB5 C
42	40.8	3.3	26338	1	V62134	HSV-2 strain SB5 C
43	40.8	3.3	117213	1	V62176	Infected cell prot
44	40.6	3.3	4257	1	V10362	The nucleotide seq
45	40.6	3.3	4257	1	V68520	

ALIGNMENTS

RESULT 1	
V49626	V49626 standard; DNA; 1245 BP.
ID	20-NOV-1998 (first entry)
AC	20-NOV-1998 (first entry)
DE	Mycobacterium tuberculosis L-Alanine dehydrogenase 40 KD antigen.
KW	ss; Alanine dehydrogenase; tuberculosis; TB; mycobacteria; L-alanine.
OS	Mycobacterium tuberculosis.
PN	W09836089-A2.
PD	20-AUG-1998.
PF	29-JAN-1998; E00483.
PR	29-JAN-1997; EP-101338.
PA	(FLOH/) FLOHE L.
PI	Flohe L, Hutter B, Kolk A, Singh M;
DR	WPI: 98-45717/39.
PT	Diagnosis of, e.g., tuberculosis from alanine dehydrogenase activity
PT	- useful for, e.g., for diagnosis, differentiation of strains,
PT	monitoring vaccination and identification of mycobacterial
PT	inhibitors.
PS	Disclosure, Fig 3.19; 55pp; German.
CC	The Mycobacterium tuberculosis L-alanine dehydrogenase (AlaDH) is used in
CC	the production of kits for diagnosing tuberculosis (TB) and other
CC	mycobacterial infections in humans or animals. Kits are used for direct
CC	diagnosis of TB on clinical samples (e.g. body fluids) and can
CC	differentiate between pathogenic and non-virulent strains, e.g. for
CC	identifying pathogens of the M. tuberculosis (M. t.) complex. The kit may
CC	also be used to identify substances that inhibit mycobacteria, for
CC	combating epidemics and for vaccination follow-up. Oligonucleotides
CC	derived from AlaDH are used similarly in diagnostic hybridisation tests,
CC	also for culture confirmation of isolated strains and for chromosome
CC	fingerprinting to detect/differentiate between mycobacteria, and for
CC	L-alanine-specific biotransformation reactions. AlaDH is an early
CC	antigen, present extracellularly after only a few days of growth, making
CC	it an ideal drug target.
SQ	Sequence 1245 BP; 238 A; 398 C; 387 G; 222 T;

Query Match 97.7%; Score 1206.4; DB 1; Length 1245;
Best Local Similarity 99.4%; Pred. No. 8.1e-242;
Matches 1234; Conservative 0; Mismatches 1; Indels 7; Gaps 2;

QY	1	ATCTTGCAGATTATCGAACATTCTTACACTGAGCTACAGTATACGAGAGAGGATATC	60
DB	1	ATCTTGCAGATTATCGAACATTCTTACACTGAGCTACAGTATACGAGAGGAGTATC	60
QY	61	ATGGCGGTGGTATTCGACGAGACCAAAACACAG-----AATTCGGGTGGCATC	114
DB	61	ATGGCGGTGGTATTCGACGAGACCAAAACACAGATTTCATTCGCGGTGGCATC	120
QY	115	ACCCGCGCGCGCTCGCGGAGTAACCCGCTGCGCATGAGGTGCTCATCGAGGAT	174
DB	121	ACCCGCGCGCGCGCTCGCGGAGTAACCCGCTGCGCATGAGGTGCTCATCGAGGAT	180
QY	175	GCCGAGAGGAGGCTGCTATCACCGAGCGGATTTCAAGCGGCGGCGGCAACTGGTC	234
DB	181	GCCGAGAGGAGGCTGCTATCACCGAGCGGATTTCAAGCGGCGGCGGCAACTGGTC	240

QY 235 GGCACCGCCGACAGGCTGTGGGCGACGCTATTATGCTCAAGGTCAAGAACCGATA 294
 DB 241 GGCACCGCCGACAGGCTGTGGGCGACGCTATTATGCTCAAGGTCAAGAACCGATA 300
 QY 295 GCGCGGAATATCGGCGCTGTGCACACGGGC -GATCTGTTCACGTTCTGCAATTGGCC 353
 DB 301 GCGCGGAATATCGGCGCTGTGCACACGGGCAGATCTGTTCACGTTCTGCAATTGGCC 360
 QY 354 GCGCTACGCTCTTGACACCATGCTGTGTGATTCGGGACCAACGCAANTGGCTACGAG 413
 DB 361 GCGTACAGCTCTTGACACCATGCTGTGTGATTCGGGACCAACGCAANTGGCTACGAG 420
 QY 414 ACCGTTCAGACCGCGGACGCGACATACCCCTGCTCCCGATGAGCGAAGTGGCCGT 473
 DB 421 ACCGTTCAGACCGCGGACGCGACATACCCCTGCTCCCGATGAGCGAAGTGGCCGT 480
 QY 474 CGACTCGCGCCGACGCTGTGGCGCTTACCCCTGATGCGAACCCAGGGGCGCGGTGTG 533
 DB 481 CGACTCGCGCCGACGCTGTGGCGCTTACCCCTGATGCGAACCCAGGGGCGCGGTGTG 540
 QY 534 CTGATGGGGGGGGGGCGCGGCTGTGAACCGGCGACGCTGTGGTGTGATGCGCGGACCC 593
 DB 541 CTGATGGGGGGGGGGCGCGGCTGTGAACCGGCGACGCTGTGGTGTGATGCGCGGACCC 600
 QY 594 GCGCGCTACAAACGACGCGCGCATGCGCAACGGCATGGGGCGGACGTTACGGTTTAGAC 653
 DB 601 GCGCGCTACAAACGACGCGCGCATGCGCAACGGCATGGGGCGGACGTTACGGTTTAGAC 660
 QY 654 ATCAACATCGCAAAACCTTGGGCACTGAGCGCGGATTCGTGGGCGCGATCCAC .ATCGC 713
 DB 661 ATCAACATCGCAAAACCTTGGGCACTGAGCGCGGATTCGTGGGCGCGATCCACACTCGC 720
 QY 714 TACTCATCGGCTTACGAGCTGTGAGGGTGGCGTCAACGCTGCGCATGCTGTGATTTGGGCG 773
 DB 721 TACTCATCGGCTTACGAGCTGTGAGGGTGGCGTCAACGCTGCGCATGCTGTGATTTGGGCG 780
 QY 774 GTCTGTGGGCGGCGGCAAGGCAACCAATTAATCTGCAATTCATCTGTGGCCATATG 833
 DB 781 GTCTGTGGGCGGCGGCAAGGCAACCAATTAATCTGCAATTCATCTGTGGCCATATG 840
 QY 834 AAACGAGTGGCTACTGTGTGATATGCAATGACGACAGGGGCGCTGTTCGAAGCTCA 893
 DB 841 AAACGAGTGGCTACTGTGTGATATGCAATGACGACAGGGGCGCTGTTCGAAGCTCA 900
 QY 894 CGACCGACCTTACGACACCGACCGGCTGGCGTGCAGACGACGCTGTTCATGGGTG 953
 DB 901 CGACCGACCTTACGACACCGACCGGCTGGCGTGCAGACGACGCTGTTCATGGGTG 960
 QY 954 GCGAATATGCGCGCTGTGCTCGGACGACGCTGACCTACGGGCTGACCAACGCGACGATG 1013
 DB 961 GCGAATATGCGCGCTGTGCTCGGACGACGCTGACCTACGGGCTGACCAACGCGACGATG 1020
 QY 1014 CCGTATGTGCTGAGCTGTGGGACCATGGCTGTGGGGGGGGGGTGGCTGCAATCCGGCA 1073
 DB 1021 CCGTATGTGCTGAGCTGTGGGACCATGGCTGTGGGGGGGGGGTGGCTGCAATCCGGCA 1080
 QY 1074 CTAGCAAGAGCTTTCGACGACGAGGAGGCTTACTGTCCGAAGGGGTGGCCACCGAC 1133
 DB 1081 CTAGCAAGAGCTTTCGACGACGAGGAGGCTTACTGTCCGAAGGGGTGGCCACCGAC 1140
 QY 1134 CTGGGGGTGCGCTTACGAGCGCGCACGCGTGTGCTGACTGTGGCGCTGTGTTAC 1193
 DB 1141 CTGGGGGTGCGCTTACGAGCGCGCACGCGTGTGCTGACTGTGGCGCTGTGTTAC 1200
 QY 1194 GCGGAGCAGCTGTGGGAGTAAGGAGGATGATGTGGGC 1235
 DB 1201 GCGGAGCAGCTGTGGGAGTAAGGAGGATGATGTGGGC 1242

RESULT 2
 ID V49510
 AC V49510 standard; DNA: 1260 BP.
 V49510;

DT 20-OCT-1998 (first entry)
 DE Mycobacterium sp. AladH DNA.
 KW Alanine dehydrogenase; AladH; ADH; diagnosis; tuberculosis; pathogen;
 KM swimmers' disease; vaccine; epidemic; infection; identification; ss.
 OS Mycobacterium sp.
 PN WO9832862-A2.
 PD 30-JUL-1998.
 PF 29-JAN-1998; E00484.
 PR 29-JAN-1997; EP-101339.
 PA (FLOH/) FLOHE L.
 PI Flohe L., Ruter B., Kolk A., Singh M;
 DR WPI: 98-427958/36.
 PT Nucleic acid encoding alanine dehydrogenase of Mycobacterium marinum
 PT - used for diagnosis of tuberculosis and other mycobacterial
 PT diseases, also for treatment and prevention, for drug screening and
 PT for bio-transformation
 PS Disclosure: Page 11: 57pp; German.
 CC This sequence encodes an alanine dehydrogenase (ADH) protein isolated
 CC from a Mycobacterium sp. This protein is used to diagnose tuberculosis
 CC and other mycobacterial infections (including 'swimmers' disease', caused
 CC by M. marinum, a fish pathogen) in humans or animals. The protein can
 CC also be used for control of epidemics and for vaccination, to screen for
 CC agents with anti-mycobacterial activity, and in bio-transformations that
 CC are specific for L-alanine. Also mycobacteria can be identified by
 CC analysis of genomic ADH sequences. ADH is an antigen that is secreted
 CC early during infection.
 SQ Sequence 1260 BP; 243 A; 403 C; 389 G; 225 T;

Query Match 97.7%; Score 1206.4; DB 1; Length 1260;
 Best Local Similarity 99.4%; Pred No. 8.1e-24;
 Matches 1234; Conservative 0; Mismatches 1; Indels 7; Gaps 2;

QY 1 ATCTTCAAGTATATGAACTTCTTACACATGAAAGCTGACATGAGAGGGTAAATC 60
 DB 16 ATCTTCAAGTATATGAACTTCTTACACATGAAAGCTGACATGAGAGGGTAAATC 75
 QY 61 ATGCGGGTGGTATTCGACCGGACCAAAACAAACG-----AATTCGGGTGGCCATC 114
 DB 76 ATGCGGGTGGTATTCGACCGGACCAAAACAAACGAAATTCGAATTCGGGTGGCCATC 135
 QY 115 ACCCGCGCGCGCTCGCGGAATTAACCGCTGTGGCATGAGGTGCTATCCAGGAGGT 174
 DB 136 ACCCGCGCGCGCTCGCGGAATTAACCGCTGTGGCATGAGGTGCTATCCAGGAGGT 195
 QY 175 GCGGAGAGGGCTCGCTATACCGACCGGATTTCAAGCGGCGGACGCGCAACTGGTC 234
 DB 196 GCGGAGAGGGCTCGCTATACCGACCGGATTTCAAGCGGCGGACGCGCAACTGGTC 255
 QY 235 GGCACCGCCGACAGGCTGTGGGCGACGCTGATTATTGCTCAAGGTCAAGAACCGATA 294
 DB 256 GGCACCGCCGACAGGCTGTGGGCGACGCTGATTATTGCTCAAGGTCAAGAACCGATA 315
 QY 295 GCGGCGGAATATCGGCGCTGTGCACACGGGC -GATCTGTTCACGTTCTGCAATTGGCC 353
 DB 316 GCGGCGGAATATCGGCGCTGTGCACACGGGCAGATCTGTTCACGTTCTGCAATTGGCC 375
 QY 354 GCGTACAGTGTTCGACCGGATGCTGTGATTCGGGACCAACGCAANTGGCTACGAG 413
 DB 376 GCGTACAGTGTTCGACCGGATGCTGTGATTCGGGACCAACGCAANTGGCTACGAG 435
 QY 414 ACCGTTCAGACCGCGGACGCGACATACCCCTGCTCCCGATGAGCGAAGTGGCCGT 473
 DB 436 ACCGTTCAGACCGCGGACGCGACATACCCCTGCTCCCGATGAGCGAAGTGGCCGT 495
 QY 474 CGACTCGCGCCGACGCTGTGGCGCTTACCCCTGATGCGAACCCAGGGGCGCGGTGTG 533
 DB 496 CGACTCGCGCCGACGCTGTGGCGCTTACCCCTGATGCGAACCCAGGGGCGCGGTGTG 555
 QY 534 CTGATGGGGGGGGGGCGCGGCTGTGAACCGGCGACGCTGTGGTGTGATGCGCGGACCC 598
 DB 556 CTGATGGGGGGGGGGCGCGGCTGTGAACCGGCGACGCTGTGGTGTGATGCGCGGACCC 615

CC derived from AladH are used similarly in diagnostic hybridisation tests, CC also for culture confirmation of isolated strains and for chromosome fingerprinting to detect/differentiate between mycobacteria, and for CC L-alanine-specific biotransformation reactions. AladH is an early CC antigen, present extracellularly after only a few days of growth, making CC an ideal drug target.

Sequence 1260 bp; 243 A; 403 C; 389 G; 225 T;

Query Match 97.7%; Score 1206.4; DB 1; Length 1260; Best Local Similarity 99.4%; Pred. No. 8.1e-242; Matches 1234; Conservative 0; Mismatches 1; Indels 7; Gaps 2;

1 ATCTTGAGATTATGAACTTCTTCACTGACGCTAAGCTATGAGAGGGTAATC 60
 16 ATCTTGAGATTATGAACTTCTTCACTGACGCTAAGCTATGAGAGGGTAATC 75
 61 ATGCGGCTGCTATTCGACCGAGACCAAAACACAG-----AATTCGGGTGGCCATC 114
 76 ATGCGGCTGCTATTCGACCGAGACCAAAACACAGATTCCATTCCGGGTGGCCATC 135
 115 ACCCGCGCGCGCTCGCGAGACTAACCCGTCGCGCATAGAGTGTCTATCCAGAGGT 174
 136 ACCCGCGCGCGCTCGCGAGACTAACCCGTCGCGCATAGAGTGTCTATCCAGAGGT 195
 175 GCGGAGAGGGCTCGGCTATACCGAGCGGATTTCAGGCGGACAGCGCAACTGTCT 234
 196 GCGGAGAGGGCTCGGCTATACCGAGCGGATTTCAGGCGGACAGCGCAACTGTCT 255
 235 GGCAGCGCGCGAGCTGAGTGGCGGACGCTGATTATTTGCTCAAGTCAAGAACGAT 294
 256 GGCAGCGCGCGAGCTGAGTGGCGGACGCTGATTATTTGCTCAAGTCAAGAACGAT 315
 295 GCGCGGAAATAGCGCGCGCGGACGACGCGGACGCTGATTATTTGCTCAAGTGGGC 353
 316 GCGCGGAAATAGCGCGCGCGGACGCTGATTATTTGCTCAAGTGGGC 375
 354 GCGTCAAGTGC.TGCAACGATGCTGTTGATTCGCGGACGCTCAATTGCTTACGAG 413
 376 GCGTCAAGTGC.TGCAACGATGCTGTTGATTCGCGGACGCTCAATTGCTTACGAG 435
 414 ACCGTCAAGCGCGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 473
 436 ACCGTCAAGCGCGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 495
 474 CGACTCGCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 533
 496 CGACTCGCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 555
 534 CTGATGGGCGGGGTGCCCGGCTGCAACCGCGGACGCTGCTGCTGCTGCTGCTGCT 593
 556 CTGATGGGCGGGGTGCCCGGCTGCAACCGCGGACGCTGCTGCTGCTGCTGCTGCT 615
 594 GCGGCTACAAAGCAGCGCGGCTGCAACCGCGGACGCTGCTGCTGCTGCTGCTGCT 653
 616 GCGGCTACAAAGCAGCGCGGCTGCAACCGCGGACGCTGCTGCTGCTGCTGCTGCT 675
 654 ATCAACATCGACAAACTTTCGCACTGCAAGCGCGGCTGCTGCTGCTGCTGCTGCT 713
 676 ATCAACATCGACAAACTTTCGCACTGCAAGCGCGGCTGCTGCTGCTGCTGCTGCT 735
 714 TACTCATCGGCTTACGAGCTGCAAGCGCGGCTGCTGCTGCTGCTGCTGCTGCT 773
 736 TACTCATCGGCTTACGAGCTGCAAGCGCGGCTGCTGCTGCTGCTGCTGCTGCT 795
 774 GTCTGTGTCAGGCGGCAAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 833
 796 GTCTGTGTCAGGCGGCAAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 855
 834 AAACCAAGTGGTACTGCTGATATAGCATGACAGGCGGCTGCTGCTGCTGCTGCT 893
 856 AAACCAAGTGGTACTGCTGATATAGCATGACAGGCGGCTGCTGCTGCTGCTGCT 915

CC derived from AladH are used similarly in diagnostic hybridisation tests, CC also for culture confirmation of isolated strains and for chromosome fingerprinting to detect/differentiate between mycobacteria, and for CC L-alanine-specific biotransformation reactions. AladH is an early CC antigen, present extracellularly after only a few days of growth, making CC an ideal drug target.

Sequence 1260 bp; 243 A; 403 C; 389 G; 225 T;

Query Match 97.7%; Score 1206.4; DB 1; Length 1260; Best Local Similarity 99.4%; Pred. No. 8.1e-242; Matches 1234; Conservative 0; Mismatches 1; Indels 7; Gaps 2;

1 ATCTTGAGATTATGAACTTCTTCACTGACGCTAAGCTATGAGAGGGTAATC 60
 16 ATCTTGAGATTATGAACTTCTTCACTGACGCTAAGCTATGAGAGGGTAATC 75
 61 ATGCGGCTGCTATTCGACCGAGACCAAAACACAG-----AATTCGGGTGGCCATC 114
 76 ATGCGGCTGCTATTCGACCGAGACCAAAACACAGATTCCATTCCGGGTGGCCATC 135
 115 ACCCGCGCGCGCTCGCGAGACTAACCCGTCGCGCATAGAGTGTCTATCCAGAGGT 174
 136 ACCCGCGCGCGCTCGCGAGACTAACCCGTCGCGCATAGAGTGTCTATCCAGAGGT 195
 175 GCGGAGAGGGCTCGGCTATACCGAGCGGATTTCAGGCGGACAGCGCAACTGTCT 234
 196 GCGGAGAGGGCTCGGCTATACCGAGCGGATTTCAGGCGGACAGCGCAACTGTCT 255
 235 GGCAGCGCGCGAGCTGAGTGGCGGACGCTGATTATTTGCTCAAGTCAAGAACGAT 294
 256 GGCAGCGCGCGAGCTGAGTGGCGGACGCTGATTATTTGCTCAAGTCAAGAACGAT 315
 295 GCGCGGAAATAGCGCGCGCGGACGACGCGGACGCTGATTATTTGCTCAAGTGGGC 353
 316 GCGCGGAAATAGCGCGCGCGGACGCTGATTATTTGCTCAAGTGGGC 375
 354 GCGTCAAGTGC.TGCAACGATGCTGTTGATTCGCGGACGCTCAATTGCTTACGAG 413
 376 GCGTCAAGTGC.TGCAACGATGCTGTTGATTCGCGGACGCTCAATTGCTTACGAG 435
 414 ACCGTCAAGCGCGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 473
 436 ACCGTCAAGCGCGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 495
 474 CGACTCGCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 533
 496 CGACTCGCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 555
 534 CTGATGGGCGGGGTGCCCGGCTGCAACCGCGGACGCTGCTGCTGCTGCTGCTGCT 593
 556 CTGATGGGCGGGGTGCCCGGCTGCAACCGCGGACGCTGCTGCTGCTGCTGCTGCT 615
 594 GCGGCTACAAAGCAGCGCGGCTGCAACCGCGGACGCTGCTGCTGCTGCTGCTGCT 653
 616 GCGGCTACAAAGCAGCGCGGCTGCAACCGCGGACGCTGCTGCTGCTGCTGCTGCT 675
 654 ATCAACATCGACAAACTTTCGCACTGCAAGCGCGGCTGCTGCTGCTGCTGCTGCT 713
 676 ATCAACATCGACAAACTTTCGCACTGCAAGCGCGGCTGCTGCTGCTGCTGCTGCT 735
 714 TACTCATCGGCTTACGAGCTGCAAGCGCGGCTGCTGCTGCTGCTGCTGCTGCT 773
 736 TACTCATCGGCTTACGAGCTGCAAGCGCGGCTGCTGCTGCTGCTGCTGCTGCT 795
 774 GTCTGTGTCAGGCGGCAAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 833
 796 GTCTGTGTCAGGCGGCAAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 855
 834 AAACCAAGTGGTACTGCTGATATAGCATGACAGGCGGCTGCTGCTGCTGCTGCT 893
 856 AAACCAAGTGGTACTGCTGATATAGCATGACAGGCGGCTGCTGCTGCTGCTGCT 915

OY	894	CGACGACACACCTTACAGACCAACCCGAGCTTGCCGTGAGACACACGGTGTTTTACACGGTG	953
Db	916	CGACGACACACCTTACAGACCAACCCGAGCTTGCCGTGAGACACACGGTGTTTTACACGGTG	975
OY	934	GCGAACATGCCCCCTCGGTGTCGAGACGCTCGACCTTACGGCGCTGACCAACCGCAGCATG	1013
Db	976	GCGAACATGCCCCCTCGGTGTCGAGACGCTCGACCTTACGGCGCTGACCAACCGCAGCATG	1033
OY	1014	CCGATATGTGTCTCGAGACTTGTGCGACATGAGTGTGGCGGGCGCGTGCAGCTGCAGATCCGGCA	1072
Db	1036	CCGATATGTGTCTCGAGACTTGTGCGACATGAGTGTGGCGGGCGCGTGCAGCTGCAGATCCGGCA	1099
OY	1074	CTAAGCCAAAGGTCTTTTGCAGCGCACGAAGGGGGCTTACTGTGCCAGACGGGTGGCCACCGAC	1133
Db	1096	CTAAGCCAAAGGTCTTTTGCAGCGCACGAAGGGGGCTTACTGTGCCAGACGGGTGGCCACCGAC	1155
OY	1134	CTGGGGGTCGCCGTTCCACCGAGCCCGCCAGCGCTCTATGGCTGATCTTCGCGCCCTGTTTAC	1193
Db	1156	CTGGGGGTCGCCGTTCCACCGAGCCCGCCAGCGCTCTATGGCTGATCTTCGCGCCCTGTTTAC	1215
OY	1194	GCGGAGCACACGTGCGGGATTAAGGGAAAGCATATATCTGGCC	1235
Db	1216	GCGGAGCACACGTGCGGGATTAAGGGAAAGCATATATGTGGCC	1257

RESULT	4
V49511	
ID	V49511 standard; DNA; 682 BP.
AC	V49511; 1998 (first entry)
DT	20-OCT-1998
DE	Mycobacterium marinum M13 DNA.
KW	Aldehyde dehydrogenase; Aldehyde; ADH; diagnosis; tuberculosis; pathogen
OS	Swimmers disease; vaccine; epidemic; infection; identification; ss.
FT	Mycobacterium marinum.
CDS	Location/Qualifiers 1..682

PN	NC09832862-A2.	/note	alanine dehydrogenase
PD	30-JUL-1998.		
PE	29-JAN-1998; E00484.		
PR	29-JAN-1997; EP-101359.		
PA	(FLOH/) FLOHE L.		
PI	Flohe L, Hutter B, Kolk A, Singh M;		
PI	MP1: 98-427958/36.		
DR	P-PSDB: W64481.		
PT	Nucleic acid encoding alanine dehydrogenase of <i>Mycobacterium marinum</i>		
PT	- used for diagnosis of tuberculosis and other mycobacterial		
PT	diseases, also for treatment and prevention, for drug screening and		
PT	for <i>in-vitro</i> transformation		
PS	Claim 1: Page 34-35; 57pp; German.		
CC	This sequence encodes an alanine dehydrogenase (ADH) protein, Mar3		
CC	isolated from <i>Mycobacterium marinum</i> . This protein is used to diagnose		
CC	tuberculosis and other mycobacterial infections (including 'swimmers		
CC	disease', caused by <i>M. marinum</i> , a fish pathogen) in humans or animals.		
CC	The protein can also be used for control of epidemics and for		
CC	vaccination, to screen for agents with anti-mycobacterial activity, and		
CC	in <i>in-vitro</i> transformations that are specific for <i>L</i> -alanine. Also mycobac		
CC	can be identified by analysis of genomic ADH sequences. ADH is an ant		
CC	that is secreted early during infection.		
CC	Sequence 682 Bp: 105 A; 254 C; 225 G; 98 T;		

Query Match	36.88;	Score 454;	DB 1;	Length 682;
Best Local Similarity	80.18;	Pred. No. 5.4e-86;		
Matches 546;	Conservative	0;	Mismatches 135;	Indels 1;
			Gaps	1;

QY 98 AATTCCGGGATGCGCCATCACCCCGGGCGCTCGGGAACTAACCCGCTGTGGCATGAGG 15
 ||||| | | ||||| ||||| | ||| | ||||| |||||
 Db 1 AATTCCGGTAAAGCATCACCCCGGGCGCTCGGGCCTTGACCAACGGCGGCACGAGG 60

158 TGTGATCCAGGAGAGTGGCCGGAGAGGGGTGGGATATACCGAGCCGATTTCAAGGGG 217
 61 TGTGATCCAGGCGGGTGGCCGGAGAGAGTTCGGCCATCTCCGAGCCGGACTTCAAGGGCG 120
 218 CAGGCGCGCACTGGTGGCACCGCCCGACCAAGTGTGGGCGCAGCTGATTTATTGCTCA 277
 121 CCGGTGTCACAGGTGATCACACCGCCGCA_CAGCG_GTGGCGGATCGGACCGTGTCTCA 180
 278 AGGTCAAAAGAACGATAGCGGGCGGAATACGCGCGCTGGAGACGGGG_GATCTGTCTCA 336
 181 AGGTCAAGAACCGATCGATCGATCGAGTAGCGCGGGCTGGCCGGGGCCAGACCTGTCTCA 240
 337 CGTCTCTGATTTGGCGCGGCTACGTGCTTGCACCGATGCGTTGTTGATTCGGACCA 396
 241 CCTACTGACACTGGCGCGCTTGGCCCGCTGCACCGATGCGCTGTGAAGTCCGGACCA 300
 397 CGTCAATTGCTACAGAGACGTCGACACCGCGGAGAGGGCACTACTCCCTGCTGTGGCCGA 456
 301 CGTCCCTCCCTACGAGAGGTTGAGACCGCGGAGGGGCAATTGGCGGTGTGTGGCGCCA 360
 457 TGAAGGAATGGCGCGGTGCACTGCGCGCGCCAGGTTGGCGCTTACCACTGATGGAACC 516
 361 TGAAGGAGTCCCGCGGGCGCTGTCCGCCCAAGCTGGGGCTTACCACCTGATGGCACCC 420
 517 AAGGGGCGCGGTTGCTGATAGGGCGGGGTGCCCGGGTGCGAACCGCGCGACGTGTGG 576
 421 ACGGGGTTGCGCGGTGCTGATAGGGCGCGCGTCCCGCGTCAAGGCTTCGCGACGTGG 480
 577 TGATGGGCGCGGACCGCGCGCTCAACGCGAGCGCGGATGGCGAAGCGCATGGGGCGGA 636
 481 TGATGGGCGCGGCGACCGCGCGATCAACGCGCGCGCGGTGGCGCAAGCGCAATGGGGGA 540
 637 CCGTTACGTTTACATCAACATCGAACAACTTGGCAACTGCAAGCGCAATGGTTCGG 696
 541 TGGTACCGGTGTGATGTCAACATCAACAAAGCTCCGCAATGAGACCGGAGTTCGGCG 600
 697 GCCGATGCACACTGCTACTCATGTAGGCGCTAGAGACTGAGAGGTGGCGTTCAAAAGTGGG 756
 601 GTGCGGTGCGGACCGCGTAC GTGAGACCTTGACACTGAGAGATGGCGCAATGCCAGCGG 660
 757 AACTGGATAGGGGCGCGTCT 778
 661 ACATGTATGCGGCGCGTCT 882

RESULT 5
N91423
ID N91423 standard; DNA: 1125 bp.
AC N91423.
DT 01-FEB-1991 (first entry)
DE Sequence of heat-resistant alanine dehydrogenase (Ah) gene with mol. wt.
DE 2 MD or less
KW Enzyme; ds.
OS *Bacillus stearothermophilus* IFO 12550.
PN J01043194.A.
PD 15-FEB-1989.
PF 10-AUG-1987; 200524.
PR 10-AUG-1987; JP-200524.
PA (NIRA) Unitika KK.
DR WPI: 89-096096/13.
PT Recombinant plasmid for transforming *Escherichia coli* -
PT obd. by connecting heat resistant alanine dehydrogenase gene to
PT vector plasmid
PR Disclosure: Fig 2, p693; 9pp; Japanese.
CC A recombinant plasmid contg. heat-resistant alanine dehydrogenase (Ah)
CC gene with mol. wt. 2 or less MD is claimed. Cells transformed with the
CC vector produce high levels of heat-resistant Ah. Transformed *E. coli*
CC is useful as a clinical inspection reagent.
SO Sequence 1125 bp; 246 A; 295 C; 367 G; 217 T;

Query Match	21.58; Score 265.2; DB 1; Length 1125;
Best Local Similarity	56.18; Pred. No. 7.1e-47;

Matches 603; Conservative 0; Mismatches 458; Indels 13; Gaps 5;

QY 61 ATGCCGCTGATTTCCGACCGAGACCAAAACAGAAATTCGGGTGGCCATCACCCG 120
 DB 1 ATGAAAGATGGGATTCCTCAAAAGAAATCAAAACATGAAGAACCGCGTCCGATCATCCG 60

QY 121 GCGCGCTGCGGAATACCCCTGCTGTCATGAGTGTCT---CATCCAGCAGGTGTC 177
 DB 61 GAGGCGGTATGACGCTGCTCAAAAGCGGGGACATGAGTGTGTAAAGAGACGAAAGC 120

QY 178 GAGAGAGGCTCGCTATACCGACCGGATTTCAAGCGCGACGCGCCGACATGTGTGCG 237
 DB 121 GCGGCTGGGTGCGGGGTTTCCGATTCGAGTATGAAGAAAGCCGGGCGACATGATGCTTC 180

QY 238 ACCGCGGACAGTGTGGGCGGACGCTGATTTATCTCAAGGTCAAGAACCGATAGCG 297
 DB 181 CGAACTGGAGAGATGCTTGGAGCGGGAGATGCTTAAAGTGAAGAGCGCGTGGCT 240

QY 298 GCGAATACGGCGCGCTGCGACAGG-GCGATCTTGTACGCTTCTTGGATTTGGCGCG 356
 DB 241 CGAGAGTTCGCTATTTTGGCCCGGATGATTTTACGATTTTGCATTTAGCCCG 300

QY 357 TCACGCTGCTGACCGATCGCTTTGATTCGCGACCAACGTCATGCTTACGAGAC 416
 DB 301 GCCGAACGCTGACGAAACCGTCTGTCGAGCAAAAGTGTGGCATCGCTTACGAGAC 360

QY 417 GTCCAGACCGCGAAGGCGGACATACCCCTGCTTGGCCCGATGAGGAATCGCCGCTGA 476
 DB 361 GTGCACTGCTGCAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420

QY 477 CTGCGCGCGGCTTGGCGCTTACCACTGATGCAACCCAAAGGCGCGCGGTGCTG 536
 DB 421 ATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480

QY 537 ATGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 596
 DB 481 CTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540

QY 597 GCGTCAACGCGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 656
 DB 541 GGGAGAAAGCGCGGCAAAATGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600

QY 657 AACATCGAACAATTCGCGACCTGACCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 716
 DB 601 AACGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660

QY 717 TCATGCGCTTACGAGCTGAGGCTGCGCAACGTCGCGACCTGCTGCTGCTGCTGCTGCTG 776
 DB 661 TCCAACTGCTACATATGCGGAGTGGCGGCAATCGGATTTGCTGCTGCTGCTGCTGCTG 720

QY 777 CTGCTGCGAGCGCGGACCGACCCAAATTAATTCGAAATTCCTGCTGCTGCTGCTGCTGCTG 836
 DB 721 TTGATCCCGGGGCGGAA--AGCGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 777

QY 837 CCAGGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 893
 DB 778 CCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 837

QY 894 CGACGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 953
 DB 838 CGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 897

QY 954 GCGAACAATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1013
 DB 898 GCGCAACATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 954

QY 1014 CCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1073
 DB 955 CCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1014

QY 1074 CTAGCAAAAGGCTTCTGCGACGACGAAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1127
 DB 1015 CTGTTAAAGGATCAACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1068

RESULT 6
 V52155/c
 ID V52155 standard; DNA: 28171 BP.

AC V52155;
 DT 23-OCT-1998 (first entry)
 DE Streptococcus pneumoniae genome fragment SEQ ID NO:22.
 KM Streptococcus pneumoniae: S. pneumoniae: genome; diagnosis; assay;
 OS computer readable medium; vaccine; pharmaceutical composition; ds.
 PN Streptococcus pneumoniae.
 PD W09818931-A2.
 PF 07-MAY-1998.
 PF 30-OCT-1997; U19588.
 PA 31-OCT-1996; US-029960.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M,
 PI Kunisch CA, Rosen CA.
 DR WPL: 98-27225/24.
 PR Computer-readable medium with recorded Streptococcus pneumoniae
 PT polynucleotide sequences - useful in diagnostic kits and assays, and
 PR pharmaceutical compositions and vaccines for Streptococcus
 PS pneumoniae.

PS Claim 1: Page 273-289, 1409pp; English.
 CC The present invention describes a computer readable medium which has
 CC the nucleotide sequences SEQ ID NO:1 to 391 (V52134 to V52524) recorded
 CC on 1t, or a representative fragment or a sequence at least 95% identical
 CC to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in SEQ ID NO:1
 CC to 391 (V52134 to V52524) are genomic fragments from Streptococcus
 CC pneumoniae. The present invention also describes an isolated nucleic acid
 CC molecule encoding a homologue of any of the fragments of the S.pneumoniae
 CC genome (SEQ ID NO:1 to 391) where the nucleic acid molecule is produced
 CC by a process comprising: (a) screening a genomic DNA library using as a
 CC probe a target sequence defined by any of the sequences in SEQ ID NO:1
 CC to 391, identifying members of the library which contain sequences
 CC that hybridize to the target sequence and isolating the nucleic acid
 CC molecules from the members; or (b) isolating mRNA, DNA or cDNA produced
 CC from an organism, amplifying nucleic acid molecules whose nucleotide
 CC sequence is homologous to amplification primers derived from the
 CC fragment of the S. pneumoniae genome to prime the amplification and
 CC isolating the amplified sequences. The computer readable medium can be
 CC used in a computer-based system for identifying fragments of the
 CC S. pneumoniae genome of commercial importance, or expression modulating
 CC fragments of the S. pneumoniae genome. Products from the present
 CC invention can be used in diagnosis kits and assays, and pharmaceutical
 CC compositions and vaccines for S. pneumoniae.
 SQ Sequence 28171 BP; 8507 A; 5080 C; 5440 G; 8144 T;

Query Match 16.5%; Score 203.6; DB 1; Length 28171;
 Best Local Similarity 54.0%; Pred. No. 5.1e-34;
 Matches 534; Conservative 0; Mismatches 434; Indels 20; Gaps 5;

QY 61 ATGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
 DB 21711 ATGTTAATCGGAATCCAAAGAAATTAATAACAAACCTGTCGCTTACACCT 21652

QY 121 GCGGCGCTGCGGCACTAACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 DB 21651 GCAAGCTTTCATGCTTGAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 21592

QY 181 GAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
 DB 21591 CTGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 21532

QY 241 GCCACCAAGGTGCGGCGGACGCTGATTTATGCTCAAGGTCAAGAAACGATAGCGCG 300
 DB 21531 GCTGTGTAAGCTTGGGCG---AGCAGAGTGTGTGTAAGAAAGAACTTAAGTCT 21475

QY 301 GAATAGCGCGCTGCGGACACGCGGCG-GATCTTGTCTACGCTTCTGCTGCTGCTGCTGCTGCTG 359
 DB 21474 GAATAGCGCTTACTTGGCGGACGATCTTCTCTCTACCTACTTGCACAGCGCGCTGCT 21415

QY 360 CGTGTGCAACCGATCGTTGGATTCCGCGACACGCTCAATTGCTACGAGACCGTC 419
 DB 21414 CCAGATATAGACATGCTATGTT-----AACAGCAAAAACCACTGAACATGTT 21367
 QY 420 CAGACCGCCGAGGCGCATACCCCTGCTTCCCGCATGAGGAAAGTCCCGGTGACCTC 479
 DB 21366 CGTGACATTCAGACACTACCGCTCTGTTCTATGAGAGGTTCAGAGTGTATG 21307
 QY 480 GCGCCGAGGTTGGCTTACACACTGATGCGAACCAAGGGGCGCGGTGTGATG 539
 DB 21306 GCGTTC/ATGAGACTACTTCTTACTAGCAAGCTGGTGGCTGTGTTCTACTT 21247
 QY 540 GCGGGGTCGCGGCTCGAACCGCGGACGTCGTGTATGGCGCGCGACCGCGGC 599
 DB 21246 GGTGTGTACAGAGTTCCTCAAAAAGAAAGTATATATGCTGTGTGTCGCT 21187
 QY 600 TACAAAGCAGCGCGCAAGCGCAACGATGGGCGACCGTTACGGTTAGACATCAAC 659
 DB 21186 ACAGATGTCGCGCGATCCCTGCTGTGTGCTCAAGTACTATTTTATATAGT 21127
 QY 660 ATGACAAACTTGGCACTCGACGCGGATCTGCGGCGC/ATCCACACTGCTACTCA 719
 DB 21126 TCCAAAGCTCTCTCACTTATAGAAAGTCTTGGAAATCAATCAACTTATGCT 21067
 QY 720 TCGGCTACGAGTGGAGGTGCGCAAGCTGCTGATGTTGGGCGCGCTCTG 779
 DB 21066 AATTCATTACATTTAGAACAAAGTGTGAGATGCTGATGTGTATGAGACATTCTC 21007
 QY 780 GTGCGAGCGCGCAAGGACCAATTAATGCTGCAATTCCTTGGCGCATGAAACCA 839
 DB 21006 ATCCCTGGTGCAGAAACACCGAATGTGACAGATGATGATCAACAAATGCTGCA 20947
 QY 840 GGTGGGCTGCTGTGATATGACATGACGAGGCGGCTGTTTGA---AGGCTACAGA 896
 DB 20946 GGTCTGTA-TGTTTACCTTGTCTGTTGACAAAGGGGCTTATGAAACACTGACCT 20888
 QY 897 CCGACCACTACGACACCGGAGTTCGCTGACGACGACGCTTCTTACTGCTGCG 956
 DB 20887 GTGACAAACGACATGAAACCGCTATATGAAACAGGGTGTCTCCATGCTGCTCC 20828
 QY 957 AACATGCGCGCTGCTGCGGAGAGCTGACCTGACCTGACCAAGCGACGATGCGG 1016
 DB 20827 AATATCCCTGCTGCTGCTGCTGCTGACCTTCAACATGCTGCTTCTCT 20768
 QY 1017 TATGTCTGAGCTTCCGACCATGCT 1044
 DB 20767 TATATGAAAGCTTGGCTGCGCAAGAT 20740

RESULT 7
 V74442/c
 ID V74442 standard; DNA: 9280 BP.
 AC V74442:
 DT 16-MAR-1999 (first entry)
 DE Staphylococcus aureus contig SEQ ID #131.
 KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
 KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KW skin infection; surgical wound infection; scalded skin syndrome;
 KW toxic shock syndrome; ds.
 KW Staphylococcus aureus.
 OS Key
 PH Location/Qualifiers
 FT misc-feature 841..900
 FT /tag-a
 FT /note- "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"
 FT misc-feature 2641..2700
 FT /tag-b
 FT /note- "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"

FT misc-feature 4441..4500
 FT /tag-c
 FT /note- "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"
 FT misc-feature 6241..6300
 FT /tag-d
 FT /note- "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"
 FT misc-feature 8041..8100
 FT /tag-e
 FT /note- "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"
 PN EP-786519-A2.
 PD 30-JUL-1997.
 PF 07-JAN-1997: 100117.
 PR 05-JAN-1996: US-009861.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA,
 PI Rosen CA.
 DR WPI: 97-374922/35.
 PT Polynucleotide(s) and proteins derived from Staphylococcus aureus -
 PT stored on computer readable medium and used in the production of
 PT anti-S.aureus vaccines
 PS Claim 1: Page 710-715: 3271pp; English.
 CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
 CC of the invention. The DNA sequences are recorded on a computer readable
 CC medium, preferably selected from a floppy or hard disk, random access
 CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
 CC the S.aureus DNA sequences allows putative functions to be assigned so
 CC that protein-encoding or regulatory regions of commercial, therapeutic or
 CC industrial importance can be obtained. Specifically, sequences which are
 CC likely to encode antigens have been identified and these polypeptides can
 CC be used in a vaccine composition against S.aureus infection. The
 CC polypeptides can also be used in a kit for the immunodetection of
 CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used
 CC for recombinant production of the polypeptides. The new DNA sequences
 CC (and their fragments) are useful as primers or probes for isolating
 CC homologues of any of the S.aureus DNA sequences contained on the
 CC computer readable medium.
 SQ Sequence 9280 BP: 2958 A: 1649 C: 1309 G: 3060 T:

Query Match 13.5%; Score 166.4; DB 1: Length 9280;
 Best Local Similarity 50.5%; Pred. No. 2,4e-26;
 Matches 483; Conservative 0; Mismatches 466; Indels 7; Gaps 3;
 QY 151 CATAGGTGCTCATCCAGGAGGCGCGGAGAGGCTCGGCTATCAGCGAGGATTC 210
 DB 9280 CATCTCTTTTATGGAAGAAATAATGCGGGTTCAGGATCATCTTTGAAAGTGTGATTAC 9221
 QY 211 AAGCGCGAGCGCGCACTGTCGACCGCGACGAGGTCGCGCGGCGCGGCTGATT 270
 DB 9220 AAGAAAGCAGGTGCTGATGATTGCTGAAACAGCAAAAGTTGG---GATGGGATAG 9164
 QY 271 TTGCTCAAGTCAAAAGACCGATAGCGCGGATACGGCGCTGCGACAGCGGC-GATC 329
 DB 9163 GTTATTAAGTCTTAAGAACCACTTGAATTCGAATATTTTAAAGAGGGCTTGA 9104
 QY 330 TTGTCAGTCTTTCATTTGGCGCGCTACGCTTGCACCGATCGTGTGATTTC 389
 DB 9103 TTATTCCTTATCTTCTATTTAGCAATGAAAGAAATTAACACAACTTGTAGATAGA 9044
 QY 390 GGCACCACTGATTCCTACAGAGCGTCCAGACCGGAAAGCGCACTACCTGCTT 449

```

Db 9043 AAGATATTAGTATTCATATATAGACTGTGCACTTACCAAGCCGCTTTACCATTTGTA 8984
Qy 450 GCCCGATGAGCGAAGTCCCGGTGACTCGCCGCCAGGTTGGCGCTTACCACTGATG 509
Db 8983 TCACCAATGAGTGTAGTACAGGAAGAAATGTGAGCTCAAGTTGGCGAGATTCCTACA 8924
Qy 510 CGAACCCAGGGGGGGCGGTGTGTATGAGCGGGGTGCGCGGGGTGCAACCGCCGAC 569
Db 8923 AACTTAAATGTTGTATGGAATTTTACTAGGTGTGTCGAGAGTACCTTAAGTAA 8864
Qy 570 GTCTGTGTATGAGCGCGCCGCGCCGCTACAAAGCAGCCGCGGATGCGCAAGCATG 629
Db 8863 GAACTATATATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8804
Qy 630 GCGCGAGCGGTATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 689
Db 8803 GGTGAGATGTATGATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 8744
Qy 690 TTCTGCGCGCGGATGCGCGCGCTACGCTACGCTACGCTACGCTACGCTACGCTACG 749
Db 8743 TTGCGGTGAGCGTGTATCATATATATATATATATATATATATATATATATATAT 8684
Qy 750 CGTCCGAGCGGTGTATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 809
Db 8683 CAAGTGTATTTAGTATATAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8624
Qy 810 TCGAATTCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 869
Db 8623 ACAGAGACATATTTAAACAAATGAAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8564
Qy 870 CAGGCGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 926
Db 8563 CAAGCGGATTTTGTAAACAACTGATAAATTTAGACATGATGATGATGATGATGATGAT 8504
Qy 927 GTGACAGACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 986
Db 8503 AAGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8444
Qy 987 ACCTACGCGCTGACCAAGCGAGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1046
Db 8443 ACGTATGCTTTAAATATATATATATATATATATATATATATATATATATATATAT 8384
Qy 1047 CGGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1102
Db 8383 AGAGAGCATTTTAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 8328

RESULT 8
T67971
ID T67971 standard: DNA: 1074 BP.
AC T67971:
DT 15-JUL-1997 (first entry)
DE H. pylori membrane protein ORF 05cp20518orf1.
KW Vaccine; prevention; treatment; infection; identification;
KW binding compound; bacterium; life cycle; activator; bacteria;
KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis;
KW membrane; amino acid; metabolism; ds.
OS Helicobacter pylori.
FH Key Location/Qualifiers
FT cds 1..1074
FT /note="no stop codon given"
PN W09640893-A1.
PD 19-DEC-1996.
PF 06-JUN-1996; U09122.
PR 07-JUN-1995; US-487032.
PR 01-APR-1996; US-630405.
PA (ASTR) ASTRA AB.
PI Berglindh OT Smith D, Mellgaard BL;
DR WPI: 97-052306/05.
PT P-SDS; W20718.
Helicobacter pylori nucleic acid sequences and related
polypeptide(s) - useful for vaccines to treat or prevent H. pylori

```

```

PT Infection, and to detect Helicobacter
PS Claim 27; Page 819; 1481bp; English.
CC The present sequence encodes a Helicobacter pylori membrane
CC protein likely to contain four membrane spanning regions.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds,
CC useful as potential H. pylori life cycle activators or inhibitors.
CC The genomic sequence of H. pylori (ATCC 55679) was determined from
CC overlapping contigs generated by mechanically shearing the
CC bacterial DNA. The sequences were analysed for ORF of at least 180
CC nucleotides, and the predicted coding regions defined by computer
CC evaluation. To identify likely H. pylori antigens for vaccine
CC development, the amino acid sequences predicted from various ORF
CC were analysed for significant homology to other known or exported
CC membrane proteins. Having identified and determined the sequences
CC of interest, particular regions can be isolated from H. pylori by
CC PCR amplification for recombinant polypeptide production, e.g. in
CC E. coli hosts.
SQ Sequence 1074 BP: 302 A; 189 C; 297 G; 286 T;

```

```

Query Match 7.2%; Score 89.2; DB 1; Length 1074;
Best Local Similarity 46.5%; Pred. No. 2e-10;
Matches 373; Conservative 0; Mismatches 413; Indels 16; Gaps 2;

```

```

Qy 265 GATTATTTGCTCAAGCTCAAGAACCATATGCGGCAATACCGCGCTGCGA-CAGCG 323
Db 133 GATTGTGGTGCATCAATGCAAGACCTTTAGCATGATACCTTTGCTCAAGAAAA 192
Qy 324 GCGATCTGTGTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 383
Db 193 GCGACTCTGTATTTATTTGATTTAGCGTATCAAAAAAGCTGTGTGAATGTATT 252
Qy 384 GATTCGCGACACAGTCAATTCCTACAGACCGTCCAGACCGCGCAAGCGCACTACCC 443
Db 253 AATAAAAAATCACTCTATTTGCACTGAACCATTCGCCGGGCTAAAGAGCATACCT 312
Qy 444 CTGCTTCCCGGATGACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 495
Db 313 ATTTTACCGCTTACAGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 372
Qy 495 -----GCTTACCACTGATGCGAACCCAGGGGGCGCGGTGTGTGTGTGTGTGTGTGT 548
Db 373 TTACTGGCTTTAGAGCATGTTAAAGGTTTATGGGTAAAGGGGTCATGCTAAGGGGTTTG 422
Qy 549 CCGGCGCTCGAACCGCGGACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 608
Db 433 TCGGTCGCGCAAGCGCTAAATGCTGTAATGAGAGCGGTGTGTGTGTGTGTGTGTGTGT 492
Qy 609 GCCGCAATCGGCAGCGCATGCGCGGCGGCGGCTTACGATCAATCAATCAATCAATCA 668
Db 493 GCGAAAGTCTTAAACCAATGAGGGCTTAAAGTAAAGATTTTAAAGATTAAGCTAGCTAAA 552
Qy 669 CTTGCGCACTCGACGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 728
Db 553 TTCAAAAACCACTTATATATATATATATATATATATATATATATATATATATATAT 612
Qy 729 GAGCTCGAGGTGCGGCAACGTCGACCTGTGTATTTGGGGCGCTCTGTGTCCAGGC 788
Db 613 AATATCAATCAAGCTTAAACGGGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 772
Qy 789 GCCAAGCGACCAATATGCTCGAATTCATCTGCTCGCATATGAAACCAAGGTGGCGTA 848
Db 673 AGCCAAACCCCTTAAGTATTAAGAAAGCATTTTAAATATACATGACAGACAAAGGGGTA 792
Qy 849 CTGTGTATATAGCATGACGACGAGGCGCTGTGTGAAGGCTCAGACGACCACTTAC 908
Db 733 GTATTTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 792
Qy 909 GACCACCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 968
Db 793 TCTAACCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 852

```


FT misc-feature 19201. 19260
 FT /*tag- k
 FT /note- "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence".
 FT misc-feature 21001. 21060
 FT /*tag- l
 FT /note- "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence".
 FT misc-feature 22801. 22860
 FT /*tag- m
 FT /note- "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence".
 FT misc-feature 24601. 24660
 FT /*tag- n
 FT /note- "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence".
 FT misc-feature 26401. 26460
 FT /*tag- o
 FT /note- "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence".
 FT misc-feature 28201. 28260
 FT /*tag- p
 FT /note- "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence".
 FT misc-feature 30001. 30060
 FT /*tag- q
 FT /note- "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence".
 EP-786519-A2.
 PD 30-JUN-1997; 100117.
 PF 07-JAN-1997; 100117.
 PR 05-JAN-1996; US-009861.
 PA (HUMAN-) HUMAN GENOME SCI INC.
 PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunesh CA,
 PI Rosen CA;
 DR WPI: 97-374922/35.
 PT Polynucleotide(s) and proteins derived from *Staphylococcus aureus*
 PT stored on computer readable medium and used in the production of
 PT anti-S. aureus vaccines
 PS Claim 1; Page 452-469; 3271pp; English.
 CC This sequence represents one of 5191 *Staphylococcus aureus* DNA sequences
 CC of the invention. The DNA sequences are recorded on a computer readable
 CC medium, preferably selected from a floppy or hard disk, random access
 CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
 CC the S. aureus DNA sequences allows putative functions to be assigned so
 CC that protein-encoding or regulatory regions of commercial, therapeutic or
 CC industrial importance can be obtained. Specifically, sequences which are
 CC likely to encode antigens have been identified and these polypeptides can
 CC be used in a vaccine composition against S. aureus infection. The
 CC polypeptides can also be used in a kit for the immunodetection of
 CC S. aureus in a sample. S. aureus is implicated in numerous human diseases,
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used
 CC for recombinant production of the polypeptides. The new DNA sequences
 CC (and their fragments) are useful as primers or probes for isolating
 CC homologues of any of the S. aureus DNA sequences contained on the
 CC computer readable medium.
 SQ Sequence 31096 BP; 11857 A; 5243 C; 5477 G; 7488 T;

Query Match 4.6%; Score 56.8; DB 1; Length 31096;
 Best Local Similarity 48.2%; Pred. No. 0.0013;
 Matches 160; Conservative 0; Mismatches 172; Indels 0; Gaps 0;
 OY 713 CTACTCATGGCCCTAGAGACTGAGGGTGGCCGTCMAACGTCGCCAGCTGTGATGGGCG 772
 DB 649 CAATCAACACCCAGAAATTTAGCAGAACAAATTAAAGAGAGATGATTTATTTCTAC 708
 OY 773 CGCTGTGGCAGCGCCGCAAGGACCCCAATTTAGTCTGATTCGATTCGCGCATAT 832
 DB 709 AATTATTTTACAGTGGCAACCGCCAAATTTGTTACTGCTGAGATGTTAATCAAT 768
 OY 833 GAACACAGTGGCTGACTGCTGATATATAGCATGACGAGCGGCTGTTGCAAGGCTC 892
 DB 769 GAAAAAGTTCAGATTAATTCATATATGATATGACAGGTGATTCATTTATGTCGT 828
 OY 893 AGACCGACACCTACGACACCCGACGTTGCCGTGACGACACGCTGTTTACTGCGT 952
 DB 829 TAGACCACTACAAATTTCTGATCCAGTGTATGAAAGAGAGGTGATTCATTTATGTCGT 888
 OY 953 GGGCAACATGCCCGCTGCTGCGGACGACGTCGACCTACGCGGTGACCAACGCGACAT 1012
 DB 889 ACCAATATCAACGAGAGAGAGTCCCAAGACCTTCACAAATGGCATTTAGCAAGAAATAT 948
 OY 1013 GCCGTATGCTGCTGAGCTTGCCGACCATGCT 1044
 DB 949 TGATTATATTATAGAAATTTGTGACAAAGGCT 980
 RESULT 11
 ID 021833/C
 AC 021833 standard; DNA: 390 BP.
 DT 08-JUN-1992 (first entry)
 DE Randomising oligonucleotide used in SPERT mRNA prepn.
 KW Systematic polypeptide e- lution by reverse translation; SPERT.
 KW Ligand binding; ss.
 OS Synthetic.
 PN W09202536-A.
 PD 20-FEB-1992.
 PF 01-AUG-1991; U05463.
 PR 02-AUG-1990; US-561968.
 PA (COLS.) UNIV OF COLORADO.
 PI Gold L, Therk C;
 DR WPI: 92-080018/10.
 PT New method of systematic polypeptide evolution by reverse
 PT translation - by linking each polypeptide in sample mixt. to
 PT individualised mRNA allowing further synthesis of selected
 PT polypeptide(s)
 PS Example, Page 55; 102pp; English.
 CC The sequence is that of an example randomising oligonucleotide which
 CC is used in the prepn. of mRNA encoding candidate polypeptides for the
 CC method of systematic polypeptide evolution by reverse translation
 CC (SPERT). The method provides a rapid way of isolating and identifying
 CC polypeptide ligands which bind to target mols. The polypeptide ligands
 CC can be used in e.g. assay methods, diagnostic procedures, cell sorting,
 CC as activators or inhibitors of enzymatic, functional, or protein, as
 CC sequencing agents, drug delivery vehicles, modifiers of hormone
 CC action and as catalysts. See also 021830-021832.
 SQ Sequence 390 BP; 125 A; 126 C; 133 G; 6 T;
 Query Match 4.0%; Score 49.2; DB 1; Length 390;
 Best Local Similarity 45.9%; Pred. No. 0.036;
 Matches 168; Conservative 0; Mismatches 198; Indels 0; Gaps 0;
 OY 425 CGCGAAGGCGCACTACCCCTGCTGCCCCGATGAGCAAGTCCGCGTACTGCGCGC 484
 DB 366 CGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 307
 OY 485 CCAGTGTGGCGCTTACCACTGATGCAACCAAGGCGCGCGTGTGCTGATGGCGG 544

Fri Jun 23 09:31:38 2000

us-09-362-485-8.rng

Page 12

Search completed: June 22, 2000, 15:10:52
Job time: 17978 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 15:13:36 ; Search time 341.15 Seconds
(without alignments)
470.360 Million cell updates/sec

Title: US-09-362-485-8
Perfect score: 1335

Sequence: 1 ATCTGAGATTAATCGAAC.....GGAGAGCATGATCGGCC 1235

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:**

1: /cgn2_6/ptodata/2/1na/5A.COMB.seq:**
2: /cgn2_6/ptodata/2/1na/5B.COMB.seq:**
3: /cgn2_6/ptodata/2/1na/5C.COMB.seq:**
4: /cgn2_6/ptodata/2/1na/5D.COMB.seq:**
5: /cgn2_6/ptodata/2/1na/6.COMB.seq:**
6: /cgn2_6/ptodata/2/1na/PCRM.COMB.seq:**
7: /cgn2_6/ptodata/2/1na/Backfiles1.seq:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45.4	3.7	833	US-08-403-852D-3	Sequence 3, Appl
2	45.4	3.7	5392	US-08-403-852D-1	Sequence 1, Appl
3	44.4	3.6	2582	US-08-816-105A-2	Sequence 2, Appl
4	43.6	3.5	28958	US-08-238-261B-6	Sequence 6, Appl
5	43.6	3.5	28958	US-08-456-837-6	Sequence 6, Appl
6	43.6	3.5	28958	US-08-457-342-6	Sequence 6, Appl
7	43.6	3.5	28958	US-08-457-646A-6	Sequence 6, Appl
8	43.6	3.5	28958	US-08-458-076A-6	Sequence 6, Appl
9	43.6	3.5	28958	US-08-764-233A-4	Sequence 6, Appl
10	43.6	3.5	28958	US-08-457-335A-6	Sequence 6, Appl
11	43.6	3.5	28958	US-08-729-214-6	Sequence 6, Appl
12	43.6	3.5	49377	US-08-764-233A-1	Sequence 6, Appl
13	42.4	3.4	1833	US-08-403-852D-6	Sequence 6, Appl
14	42.4	3.4	2185	US-08-173-308-3	Sequence 3, Appl
15	42.4	3.4	2185	US-08-165-310-3	Sequence 3, Appl
16	41.6	3.4	1620	US-08-461-775-10	Sequence 10, Appl
17	41.6	3.3	20235	US-07-642-734C-3	Sequence 3, Appl
18	40.6	3.3	4257	US-08-690-473-1	Sequence 1, Appl
19	40.6	3.3	12001	US-08-458-568A-11	Sequence 1, Appl
20	39.6	3.2	2414	5248599-1	Patent No. 5248599
21	39.6	3.2	2668	US-08-461-775-11	Sequence 11, Appl
22	39.4	3.2	459	US-08-387-942C-35	Sequence 35, Appl
23	39.4	3.2	1978	5212296-8	Patent No. 5212296
24	39.4	3.2	12588	US-08-387-942C-1	Sequence 1, Appl
25	39.2	3.2	420	US-08-470-179-148	Sequence 148, App
26	39.2	3.2	44377	US-08-804-227C-7	Sequence 7, Appl
27	39.2	3.2	44377	US-08-804-198-1	Sequence 1, Appl

28	39	3.2	43280	US-08-804-227C-1	Sequence 1, Appl
29	38.8	3.1	8051	US-08-576-626A-2	Sequence 2, Appl
30	38.8	3.1	11219	US-07-642-734C-1	Sequence 1, Appl
31	38.6	3.1	2064	US-08-343-428-1	Sequence 1, Appl
32	38.6	3.1	11219	US-07-642-734C-1	Sequence 1, Appl
33	38.4	3.1	1215	US-08-947-726A-1	Sequence 1, Appl
34	38.4	3.1	2048	US-08-776-251-1	Sequence 1, Appl
35	38.2	3.1	2109	US-08-555-568B-20	Sequence 20, Appl
36	38	3.1	474	US-08-403-852D-14	Sequence 14, Appl
37	38	3.1	30001	US-08-125-468-1	Sequence 1, Appl
38	38	3.1	30001	US-08-474-933-1	Sequence 1, Appl
39	37.8	3.1	3252	US-08-809-740A-1	Sequence 1, Appl
40	37.8	3.1	3252	US-08-809-740A-4	Sequence 4, Appl
41	37.8	3.1	15664	US-08-402-282-3	Sequence 3, Appl
42	37.8	3.1	15664	US-08-508-004-3	Sequence 3, Appl
43	37.8	3.1	15664	US-08-402-056-3	Sequence 3, Appl
44	37.8	3.1	15664	US-08-402-068-3	Sequence 3, Appl
45	37.6	3.0	1524	5512669-1	Patent No. 5512669

ALIGNMENTS

RESULT 1
US-08-403-852D-3
Sequence 3, Application US/08403852D
Patent No. 5891695
GENERAL INFORMATION:
APPLICANT: Blanc, Veronique
APPLICANT: Blanche, Francis
APPLICANT: Crouzet, Joel
APPLICANT: Jacques, Nathalie
APPLICANT: Lacroix, Patricia
APPLICANT: Thibaut, Denis
APPLICANT: Zagorec, Monique
APPLICANT: Debussche, Laurent
TITLE OF INVENTION: De Crecy-Lagard, Valerie
TITLE OF INVENTION: Polypeptides Involved In The
TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences
TITLE OF INVENTION: Coding For These Polypeptides And Their Use
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flinegan, Henderson, Farabow, Garrett & Dinner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/403,852D
FILING DATE: 10-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR 93/00923
FILING DATE: 25-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/11441
FILING DATE: 25-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03806.0054-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 833 base pairs
TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: S.pristinaespiralis
FEATURE:
NAME/KEY: CDS
LOCATION: 1..833
US-08-403-852D-3

Query Match
Best Local Similarity 48.0%; Score 45.4; DB 3; Length 833;
Matches 130; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 542 CGGGGTCGGGGCGCTGCGAAGCGGGCGAGCTGCTGATCGGCGCGCCGACCGCGGCTA 601
DB 327 CGAGGGCCGACCGGGGAGACTGATGAGCTGCTCGCGGCTGTGGAGAGCTTGCACGA 386
QY 602 CAAGCGAGCCCGCATCGCCAGCGGATGGCGCGACCGCTTACGTTTACATCAACAT 661
DB 387 CGACGCTTCTGTCACAGACGGCGCGAGCGCTGTACTGGCGGTGCCGCCGCTCACCA 446
QY 662 CGACAACTTCGCAACTGACGCGGATGTCGGCGCGGATCCACACTCGTACTATC 721
DB 447 ACTGACCGACCGAGCGGAGGAGCTGAGCTGGCGCCCTCAAGCTGCGCCGCC 506
QY 722 GGCCTACGAGCTCGAGGGTGGCGCTCAAGCTGAGCTGTGATGGGGCGCTCTGT 781
DB 507 GCAGGGCCGACCGCGTGTGTCGCTGACGGCGCGCCGCTGCGCGCGCGGAGCTTGT 566
QY 782 GCCAGGGCGCAGGCGACCAATTAGTCTCG 812
DB 567 CTGCTCGACGAGCGCGCGACCGCGCTCG 597

RESULT 2

US-08-403-852D-1

Sequence 1, Application US/08403852D

Patent No. 5891695

GENERAL INFORMATION:

APPLICANT: Blanc, Veronique

APPLICANT: Blanchet, Francis

APPLICANT: Crouzet, Joel

APPLICANT: Jacques, Nathalie

APPLICANT: Lacroix, Patricia

APPLICANT: Thibaut, Denis

APPLICANT: Zagorec, Monique

APPLICANT: Debussche, Laurent

APPLICANT: De Crecy-Lagard, Valerie

TITLE OF INVENTION: Polypeptides Involved In The

TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences

TITLE OF INVENTION: Coding For These Polypeptides And Their Use

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: Flunegan, Henderson, Farabow, Garrett & Dunner

STREET: 1300 I Street, N.W., Suite 700

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/403,852D

FILING DATE: 10-MAY-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/FR 93/00923

FILING DATE: 25-SEP-1993
PRIOR APPLICATION DATA: FR 92/11441
APPLICATION NUMBER: FR 92/11441
FILING DATE: 25-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03806.0054-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5392 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: S.pristinaespiralis
US-08-403-852D-1

Query Match
Best Local Similarity 48.0%; Score 45.4; DB 3; Length 5392;
Matches 130; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 542 CGGGGTCGGGGCGCTGCGAAGCGGGCGAGCTGCTGATCGGCGCGCCGACCGCGGCTA 601
DB 3018 CGAGGGCCGACCGGGGAGACTGATGAGCTGCTCGCGGCTGTGGAGAGCTTGCACGA 3077
QY 602 CAAGCGAGCCCGCATCGCCAGCGGATGGCGCGACCGCTTACGTTTACATCAACAT 661
DB 3078 CGACGCTTCTGTCACAGACGGCGCGAGCGCTGTACTGGCGGTGCCGCCGCTCACCA 3137
QY 662 CGACAACTTCGCAACTGACGCGGATGTCGGCGCGGATCCACACTCGTACTATC 721
DB 3138 ACTGACCGACCGAGCGGAGGAGCTGAGCTGGCGCCCTCAAGCTGCGCCGCC 3197
QY 722 GGCCTACGAGCTCGAGGGTGGCGCTCAAGCTGAGCTGTGATGGGGCGCTCTGT 781
DB 3198 GCAGGGCCGACCGCGTGTGTCGCTGACGGCGCGCCGCTGCGCGCGGAGCTTGT 3257
QY 782 GCCAGGGCGCAGGCGACCAATTAGTCTCG 812
DB 3258 CTGCTCGACGAGCGCGCGACCGCGCTCG 1288

RESULT 3

US-08-816-105A-2

Sequence 2, Application US/08816105A

Patent No. 5989882

GENERAL INFORMATION:

APPLICANT: Gruener, Anneliese; Dellweg, Hans-Georg; Lenz,

APPLICANT: Jurgen; Schroder, Werner; Pape, Hermann;

APPLICANT: Goebel, Klaus; Schaper, Beate; Hemker, Michael;

APPLICANT: Piepersberg, Wolfgang; Distler, Jurgen;

APPLICANT: Straumann, Ansgar

TITLE OF INVENTION: PROCESSES FOR PREPARING ACARVOSYL

TITLE OF INVENTION: TRANSFERASE AND FOR USING IT IN THE

TITLE OF INVENTION: CONVERSION OF ACARBOSE HOMOLOGUES

TITLE OF INVENTION: INTO ACARBOSE, FOR THE PREPARATION

TITLE OF INVENTION: OF ACARBOSE HOMOLOGUES

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sprung Kramer Schaefer & Briscoe

STREET: 660 White Plains Road

CITY: Tarrytown

STATE: New York

COUNTRY: USA

ZIP: 10591-5144

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,342
FILING DATE: 01-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHEICAL: NO
ANTI-SENSE: NO
US-08-457-342-6

Query Match 3.5%; Score 43.6; DB 1; Length 28958;
Best Local Similarity 44.8%; Pred. No. 0.097; Mismatches 354; Indels 4; Gaps 3;
Matches 290; Conservative 0;
QY 148 GGCATGAGGTGCTTCATCAGGAGGTGCGGAGGCTTGCCTATCAGCCGAGCGGAT 207
DB 25306 GGCACGACGACCTACCTCGGAGATGAGCTATTCCTGCGACACACCGCGCCAC 25365
QY 208 TTCAGAGGCGGAGCGGCGCACTGCTGCGACCGCGGACCAAGTGTGGCCGACGCTAT 267
DB 25366 TTCGACGACCGCGGCGCTCTCGAGCCCAACCGGACGAGAGCTCTCTCGCGGCTCGAC 25425
QY 268 TTATTCCTCAAGTCAAA-AGAACCGATAGCGGCGGAGATACGCGCGCTCGACACGCGG 326
DB 25426 TGGCTGGCCAGAGACAGCGCGCGCGGACGACCGTCTCGGACGAGCGGAAGCCAGCG 25485
QY 327 ATCTTGTACGCTTCTTGCATTTGCGCGCTCACGCTGCTGCACCGATCGTTGTTGAT 386
DB 25486 AAGCTGCTGCTTCTTCTCTGCGGCAAGGCTGCAAGTGGGAAGGATGCGCTCTCGCTG 25545
QY 387 TCCGACACGATCAATTCCTACGAGACCGGTCACACCGCGGACGACGACCTACCCCTG 446
DB 25546 CTCGACTCTCTCGCGCTCTTCGCGGCTCAGCTGGAAGCATGCGAGCGCGCTCGCTCT 25605
QY 447 CTGGCCCGATAGCGAGATGCGCGGCTGCACTGCGCGGACGAGTTGGCCTTACGAC--C 504
DB 25606 CACGTCGAGTGAAGCTGCTGCGCGCTGCGCGCGGACGAGGCGCGCGCTCTCGAC 25665
QY 505 TATGCGAACCAAGGCGCGCGGCTGCTGATGAGGCGGCGCGCGCGCTGCAACCG 564
DB 25666 CCGCTGAGCTGACAGCGCGCGCTCTTTCGCGCTATGCTCTCTCGCGCGCTCTG 25725
QY 565 CCGAGC-TCGTGATGATCGCGCGCGCGGACCGCGCTACACGACCGCATCGCAAC 623
DB 25726 CGCTGCGTGGGTAGAGCGCGCGCGGCTGCGCGGACAGTACAGGCGAGATCGCGCG 25785
QY 624 GGCATGAGGCGGACGCTTACGTTTCAACATCAACATGACAACACTTGGCAACTCGAC 683
DB 25786 GCGTCTGCTGAGGCGGCTCTCTCTCGAGGAGCGCGCGCGCGCTCTGCGAGC 25845
QY 684 GCCGATCTGCGCGCGGATTCACACTGCTACTATGCGGCTAGAGGCTGAGGCTGCG 743

DB 25846 AAGCGCTCACCACCGTCCCGGCAACGGGCGCATGCGCGCTGAGTCTCGCGCTCC 25905
QY 744 GTCAAACGTCGAGCTGATTTGGGCGGCTTCGCTGCTGCGAGCGCC 791
DB 25906 GACCTTCAGACCTTACTCTCTCTGCGGCGAGAGGCTCTCATGCGC 25953

RESULT 7
US-08-457-646A-6
Sequence 6, Application US/08457646A
Patent No. 5679550
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Gaffney, Thomas Deane
APPLICANT: Lam, Stephen Ting
APPLICANT: Hammer, Phillip E.
APPLICANT: Unnes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,646A
FILING DATE: 01-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHEICAL: NO
ANTI-SENSE: NO
US-08-457-646A-6
Query Match 3.5%; Score 43.6; DB 1; Length 28958;
Best Local Similarity 44.8%; Pred. No. 0.097; Mismatches 354; Indels 4; Gaps 3;
Matches 290; Conservative 0;
QY 148 GGCATGAGGTGCTTCATCAGGAGGTGCGGAGGCTTGCCTATCAGCCGAGCGGAT 207
DB 25306 GGCACGACGACCTACCTCGGAGATGAGCTATTCCTGCGACACACCGCGCCAC 25365
QY 208 TTCAGAGGCGGAGCGGCGCACTGCTGCGACCGCGGACCAAGTGTGGCCGACGCTAT 267

Db 25366 TTCGAGCACCAGGCGGCTCTGTGAGCCACCAACCGCGAGCTCTCTCCGCGCTGAC 25425
Qy 268 TTATGCTCAAGGTCA-AAGAACGATAGCGCGGAATACGCGCGCTCGACACAGCGCG 326
Db 25426 TCGCTCGCGCCAGGAGACCGCGCCGAGCACCGTCTCGAGGAGGAGGAGCGACCGC 25485
Qy 327 ATCTGTTCACGTTCTTGTGATTTGGCCCGCTGACCTGCTTGCACCGATGCTTGTGAT 386
Db 25486 AAGCTCGTCTTCTGCTTCTCTGAGGAGGCTGAGTGGGAGGAGGAGGAGGAGGAGG 25545
Qy 387 TCCGCGACACGATCAATGAGCTAGAGACCGCTCAGACCGCGCGCGACACTACCTG 446
Db 25546 CTCGACTCTCGCGCGCTCTTCCGCGCTCAGCTCGAAGATGAGAGCGCGCGCTGCT 25605
Qy 447 CTGCGCGCGATGAGGAGGAGGCGCGGCTGAGCTCGCGCGCGCGCGCGCTTACAC--C 504
Db 25606 CACGTCGAGTGGAGCGCTCTGCTCTGCGCGCGCGCGCGCGCGCGCTCTGAC 25665
Qy 505 TGATCGGAGAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 564
Db 25666 CGCGTCGAGCTGTACAGCGCGCGCTCTTTCGCGCTGATGCTCTCTGCGCGCTG 25725
Qy 565 CCGAGC-TGCTGTGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 623
Db 25726 CGCTCGCTCGCGGTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 25785
Qy 624 GGCATGCG 683
Db 25786 GCGTCTGCTGCG 25845
Qy 684 GCGGAGTCTGCG 743
Db 25846 AAGCGCTACACCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 25905
Qy 744 GTCAACGTCG 791
Db 25906 GACCTCCAGACTACCTGCTCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 25953

RESULT 8
US-08-458-076A-6
Sequence 6, Application US/08458076A
Patent No. 5698425

GENERAL INFORMATION:

APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Gafney, Thomas Deane
APPLICANT: Lam, Stephen Ting
APPLICANT: Hammer, Phillip E.
TITLE OF INVENTION: Genes for the synthesis of
TITLE OF INVENTION: antipathogenic substances
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 St. Line Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458.076A
FILING DATE: 01-JUN-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-Jun-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-458-076A-6

Query Match 3.5%; Score 43.6; DB 2; Length 28958;
Best Local Similarity 44.8%; Pred. No. 0.097;
Matches 290; Conservative 0; Mismatches 354; Indels 4; Gaps 3;

Qy 148 GGCATGAGGTGCTCATTCACGAGCGAGTGCAGAGGCTCGCTATCACGAGCGGAT 207
Db 25306 GCCCAGCAGACCTCACCCCTCGGGATGTGGCTATTCGCTGCGCACACCGCGCGCC 25365
Qy 208 TTCAAG: YGCGAGGCGGGCAACTGTGTGAGCAGCGCGCGCGCGCGCGCGCGCGCG 267
Db 25366 TTGAGC: YCGCGCGCGCTCTCTGAGCCACACCGGAGAGCTCTCTCTCGCGCTGAC 25425
Qy 268 TTATGCTCAAGGTCA-AAGAACGATAGCGCGGAATACGCGCGCTGCGAGCGGCG 326
Db 25426 TCGCTCGCGCCAGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 25485
Qy 327 ATCTGTTCACGTTCTTGTGATTTGGCCCGCTGACCTGCTTGCACCGATGCTTGTGAT 386
Db 25486 AAGCTCGTCTTCTGCTTCTCTGAGGAGGCTCGAGTGGAGGAGGAGGAGGAGGAGG 25545
Qy 387 TCCGCGACACGCTAATGCGCTACGAGCGCTCAGACCGCGCGCGCGCGCGCGCGCGCG 446
Db 25546 CTCGACTCTGCGCGCGCTCTTCCGCGCTCAGCTCGAAGCATGAGCGCGCGCTGCT 25605
Qy 447 CTGCGCGCGATGAGCGAGTGCAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 504
Db 25606 CAGTTCGAGTGGAGCGCTGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 25665
Qy 505 TGATCGGAGAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 564
Db 25666 CGCGTCGAGCTGAGCG 25725
Qy 565 CCGAGC-TGCTGTGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 623
Db 25726 CGCTCGCTCGCGGTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 25785
Qy 624 GGCATGCG 683
Db 25786 GCGTCTGCTGCGCGCGCTCTTCCGCGCTCAGCTCGAAGCATGAGCGCGCGCGCG 25845
Qy 684 GCGGAGTCTGCG 743
Db 25846 AAGCGCTACACCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 25905
Qy 744 GTCAACGTCG 791
Db 25906 GACCTCCAGACTACCTGCTCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 25953


```

REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-457-335A-6

```

```

Query Match          3.5%; Score 43.6; DB 2; Length 28958;
Best Local Similarity 44.8%; Pred. No. 0.097;
Matches 290; Conservative 0; Mismatches 354; Indels 4; Gaps 3;

```

```

OY 148 GGCATGAGTGTCTCATCCAGGAGTGGCGGAGAGGGCTCGCTATCCAGCGGGAT 207
DB 25306 GCGCCAGAGAGCTTACCTCCGCGATGCGCTTTCGCTGCGCACCCGCCCHC 25365
OY 208 TTCAAGGGGCGAGCGCGCACTGTGCGCACCGCCAGCAGTGTGGCGCGCTGAT 267
DB 25366 TTGAGAACCGGGCGCTCTGTAGCCCAACCGGAGAGAGCTCTCCGCGCTGAC 25425
OY 268 TTATGCTCAAGTCAA-AGAACCATAGGGGCGGATACGGCCCGCTGCGACAGGGCG 326
DB 25426 TCGCTCGCCAGGACAGCCCGCCCGGACACCGTCTCGAGGAGCGAAGCCACGCG 25485
OY 327 ATCTGTTCACGTTCTTCGATTTGGCCGCGCTGACGTGCTTGACGATGCTTGGAT 386
DB 25486 AAGCTCGTCTGCTTCTTCTTCTGCGCAAGGCTGCGAGTGGGAAGGATGGCCCTGCGT 25545
OY 387 TCCGGCACACGCTCAATGCTCTAGAGACCGTCCAGACCGCGGAAGGCGACTACCCCTG 446
DB 25546 CTGCACTCTCTCCCGCTCTTCCGCTCAGTCAAGCATGCGAGCGCGCTCTCTCT 25605
OY 447 CTTGGCCCGATGAGGAGTCCCGGCTGACTCGCCCGCCAGTTGGGCTTACGAC--C 504
DB 25606 CACGTCGAGTGAAGCTCTGCTCCGCTCGCCCGCGAGAGGGGCGCCCTCTCTGAC 25665
OY 505 TATATCGAAGCCCAAGGGGCGCGGCTGTGCTGATGAGCGGGGCTGCGCGCTGCAACCG 564
DB 25666 CCGCTCGAGCTGTGACAGCCCGCTCTTGGCGTATGATGCTCTCTGCGGCGCTCTG 25725
OY 565 CCGAGC-TGCTGTGATGCGCGCGCGGACCGCGCTACACGACCGCGCATGGCCAAC 623
DB 25726 CGCTGCTGCGGCTGAGAGCCCGCGCGCTGCGGCGCATGAGGGCGAGATCGCGCG 25785
OY 624 GGCATGGGGGCGAGGCTTCTAGACATCAACATGCAAACTTGGGAACCTGAC 683
DB 25786 GCGTTCGTGCGAGGCGCTCTCTCTCGAGAGCGGGCGCGCATGCGCGCTGTCGAC 25845
OY 684 GCGGATTTGCGCGCGGATCCACATCGCTACTCATCGGCTAGAGCTGCAATGTCG 743
DB 25846 AAAGGCGCTACACACCGTGGCGCGCAAGGGGCGCATGGCGCGCTGTGAGAGTCTG 25905
OY 744 GTCAAAAGTGGCGAGCTGTGATTTGGGGCGCGCTGTGTCG 366GCC 791
DB 25906 GACCTCCAGAGCTTACCTGCTCTCTGCGGCGCGAGAGGCTCTCATGCGC 25953

```

```

RESULT 11
US-08-729-214-6
Sequence 6, Application US/08729214
Patent No. 5817502
GENERAL INFORMATION:
APPLICANT: Ligon, James M.
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew

```

```

APPLICANT: Hammer, Phillip E.
APPLICANT: van Pee, Karl-Heinz
APPLICANT: Kirner, Sabine
TITLE OF INVENTION: Genes for the synthesis of
TITLE OF INVENTION: antipathogenic substances
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESS: Ciba-Geigy Corporation
STREET: 520 White Plains Road
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/729,214
FILING DATE: TBA
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Melis, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1506/CIP5
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-729-214-6

```

```

Query Match          3.5%; Score 43.6; DB 2; Length 28958;
Best Local Similarity 44.8%; Pred. No. 0.097;
Matches 290; Conservative 0; Mismatches 354; Indels 4; Gaps 3;

```

```

OY 148 GGCATGAGTGTCTCATCCAGGAGTGGCGGAGAGGGCTCGCTATCCAGCGGGAT 207
DB 25306 GCGCCAGAGAGCTTACCTCCGCGATGCGCTTTCGCTGCGCACCCGCCCHC 25365
OY 208 TTCAAGGGGCGAGCGCGCACTGTGCGCACCGCCAGCAGTGTGGCGCGCTGAT 267
DB 25366 TTGAGAACCGGGCGCTCTGTAGCCCAACCGGAGAGAGCTCTCCGCGCTGAC 25425
OY 268 TTATGCTCAAGTCAA-AGAACCATAGGGGCGGATACGGCCCGCTGCGACAGGGCG 326
DB 25426 TCGCTCGCCAGGACAGCCCGCCCGGACACCGTCTCGAGGAGCGAAGCCACGCG 25485
OY 327 ATCTGTTCACGTTCTTCGATTTGGCCGCGCTGACGTGCTTGACGATGCTTGGAT 386
DB 25486 AAG-TGCTGTGCTTCTTCTGCGGCGCAAGGCTCGAGTGGGAAGGATGGCCCTTCTG 25545
OY 387 TCCGACACACGCTCAATGCTCTACGAGACCGTCCAGACCGCGGAAGCGGACGCTG 446
DB 25546 CTGCACTCTCTCCCGCTTCTCGGCTCAGTGTGAGATGCGAGATGCGAGCGCGCTGCT 25605
OY 447 CTTGGCCCGATGAGGAGTCCCGGCTGACTCGCGCCAGAGTGTGGCGCTTACAC--C 504
DB 25606 CACGTCGAGTGAAGCTGTGCTCTCTGCGCGCGAGCGGAGGCGCGCCCTCTCTGAC 25665
OY 505 TATATCGAAGCCCAAGGGGCGCGGCTGTGCTGATGAGGCGGGGCTCCCGGCTGCAACCG 564
DB 25666 CCGCTCGAGTGTACAGCCCGCGCTCTTGGCGCTCATGATGCTCTCTGCGGCGCTCTG 25725

```

```

0Y 565 CCGAGC-TCGTGTGATATGGGGCGGGGACCGCGGCTACCAAGCAGCCCGATCGCAAC 623
Db 25726 CGCTGCTGTGGCGTGAAGCCCGCCCGCTGTCTGGGCACTGACAGGGCGAGATCGCCG 25785
0Y 624 GGCAATGGGCGCGACCGTTACGGTTCTAGACATCAACATGACAAATTCGGCACTGCAC 683
Db 25786 GCGTTCGTGTGAGGGGCTCTCTCCCTCGAGGAGAGGGGGCCGATCGCGCCCTGGGAC 25845
0Y 684 GCCGAGTTCGTGGGGCGCGGATCCACACTGCGTCTACTATCGGCTTAGAGCTCGAGGTGC 743
Db 25846 AAGGCGCTACACACCGTCGCGCGGCAAGGGGCGCATGCGCGCGTGTGAGCTCGGCGCTTC 25905
0Y 744 GTCGAACGTCGCGACCTGTGATGTGGGCGCGCTCTGTGTCAGCGCC 791
Db 25906 GACCTTCAGACTACTCTGCTCTCCCTGGGGGAGAGGCTCTCATGCCC 25953

RESULT 12
US-08-764-233A-1
Sequence 1, Application US/08764233A
Patent No. 5716849
GENERAL INFORMATION:
APPLICANT: Ligon, James M.
APPLICANT: Schupp, Thomas
APPLICANT: Beck, James J.
APPLICANT: Hill, Dwight S.
APPLICANT: Neff, Snezana
APPLICANT: Reals, John A.
TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSER: Ciba-Geigy Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/764,233A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/729,214
FILING DATE: 09-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: 1506/CI16
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 49377 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Sorangium cellulosum
IMMEDIATE SOURCE:
CLONE: p98/1, pJ13, and pVK15
FEATURE:
NAME/KEY: misc_feature
LOCATION: 383..760

```

OTHER INFORMATION: /pr duct= "SorB"
OTHER INFORMATION: /note= "This gene encodes a protein that is highly homologous to type I PKS such as eryA from Saccharopolyspora erythraea."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 927..19874
OTHER INFORMATION: /product= "SorA"
OTHER INFORMATION: /note= "Gene product is highly homologous to type I PKS t
OTHER INFORMATION: are known to be involved in the synthesis of polycetide
OTHER INFORMATION: compounds."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 942..7115
OTHER INFORMATION: /product= "Module 1 of SorA"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7203..12884
OTHER INFORMATION: /product= "Module 2 of SorA"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 13455..19616
OTHER INFORMATION: /product= "Module 3 of SorA"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19871..46318
OTHER INFORMATION: /product= "SorB"
OTHER INFORMATION: /note= "Gene product is highly homologous to type I PKS ge
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19870..24556
OTHER INFORMATION: /product= "Module 1 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 24638..30820
OTHER INFORMATION: /product= "Module 2 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 30881..35446
OTHER INFORMATION: /product= "Module 3 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 35528..40114
OTHER INFORMATION: /product= "Module 4 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 40190..46318
OTHER INFORMATION: /product= "Module 5 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 46851..47891
OTHER INFORMATION: /product= "SorM"
OTHER INFORMATION: /note= "The protein encoded by the sorm gene is highly
OTHER INFORMATION: homologous to the methyltransferase from Streptomyces
OTHER INFORMATION: hygroscopicus that is involved in the synthesis of the
OTHER INFORMATION: polyketide rapamycin."
US-08-764-233A-1

Query Match 3.5%; Score 43.6; DB 2; Length 49377;
Best Local Similarity 44.8%; Pred. No. 0.11; Mismatches 354; Indels 4; Gaps 3;
Matches 290; Conservative 0;

OY 148 GCCATAGTGTCTCATTCAGGCAGTGCCGCGAGAGGGCTCGGTACTACCGACGCCGAT 207
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 41609 GCCCAGCACGACCTACCCTCGCGGATGTGCGCTATTGCTGGCCACCAACCGCGCCAC 41668
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 208 TTCAAAGGCGGACGAGGCGCAACTGGTGGCGAACCGCGACCAAGCTGTGGCGCGAGCTGAT 267
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 41669 TTTCGACACCGCGCCGCTCTGTGTAGCCACAACCGCGAGACTCTCTTCGCGCTGAC 41728
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 268 TTAATGCTCAAGGTCGAAGAACCGATAGCGCGGAGATAACGGCCCGCTCGACACGGGCG 326

```

DB 41729 TCGCTGCGCCAGGACAGCCGCCGCCGAGCAGCCTGCTCGAGCGAGGAGGAGCCAGCGC 41788
QY 327 ATCTTGTACGCTCTTCTGCAATTTGGCCGCGCTACAGTGCCTTGCACCGATCCGTTGTGAT 386
DB 41789 AAGCTGCTGCTTCTTCTTCTTCTGCGCAAGGCTCGCAGTGGGAAGATGGCCCTCTCGCTG 41848
QY 387 TCGGCGACCGATGCAATTCCTTACAGACGCTCGACAGCCGCGCAAGGCGCAGCTACCCCTG 446
DB 41849 CTCGACTCTTCCGCCCTCTTCCGCGCTCGAGCTCGAAGCATGGAGCGCGCGCTCGCTCT 41908
QY 447 CTTCGCCGATGAGCAGAGTCCGCGGCTGCACTGCGCCGAGGTTGGCGCTTACAC -C 504
DB 41909 CAGCTGAGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 41968
QY 505 TGTGCGAAGCCAGAGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 564
DB 41969 CCGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 42028
QY 565 CCGAGC-TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 623
DB 42029 CGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 42088
QY 624 GGCATGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 683
DB 42089 GCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 42148
QY 684 GCGGAGTTTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 743
DB 42149 AAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 42208
QY 744 GTCGAAAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 791
DB 42209 GACCTGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 42256

```

RESULT 13

US-08-403-852D-6
Sequence 6, Application US/08403852D

Patent No. 5891695

GENERAL INFORMATION:

APPLICANT: Blanc, Veronique
APPLICANT: Blanc, Francis
APPLICANT: Crozet, Joel
APPLICANT: Jacques, Nathalie
APPLICANT: Lacroix, Patricia
APPLICANT: Thibaut, Denis
APPLICANT: Zagorec, Monique
APPLICANT: Debussche, Laurent
TITLE OF INVENTION: Polypeptides Involved In The
TITLE OF INVENTION: Biosynthesis of Streptogramins, Nucleotide Sequences,
TITLE OF INVENTION: Coding For These Polypeptides And Their Use
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/403,852D
FILING DATE: 10-MAY-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/FR 93/00923
FILING DATE: 25-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/11441

```

FILING DATE: 25-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03806.0054-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1833 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: S.pristinaespiralis
FEATURE:
NAME/KEY: CDS
LOCATION: 103..1689
US-08-403-852D-6

```

Query Match 3.4% Score 42.4; DB 3: Length 1833;
Best Local Similarity 50.5%; Pred. No. 0.088;
Matches 103; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

```

QY 532 TCGCTATGGCGGCGGTGCGCGCGCTGCAACCGCGCAGCTGCTGCTGCTGCTGCTGCTGCTG 591
DB 413 TGATGCTGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 472
QY 592 CCGCGGCTACAGCGCAGCCGCGATGCGCAAGGCGATGCGCGCGCGCTTACGCTTAG 651
DB 473 CCGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 532
QY 652 ACATGACATGCAACAAGTGGCAACTGCAAGCGCGAGTTGCGCGCGCGATTCGCGCGCGCACTG 711
DB 533 CCACCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 592
QY 712 GCTACTATGCGCGCTGAGAGCTGCG 735
DB 593 TGTGGAGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 616

```

RESULT 14

US-08-173-508-3
Sequence 3, Application US/08173508

Patent No. 5616485

GENERAL INFORMATION:

APPLICANT: Bartfeld, Daniel
APPLICANT: Butler, Michael J.
APPLICANT: Hadary, Dany
APPLICANT: Jenish, David
APPLICANT: Krieger, Timothy
TITLE OF INVENTION: STREPTOMYCIN PROTEASES AND IMPROVED
TITLE OF INVENTION: STREPTOMYCIN STRAINS FOR EXPRESSION OF PEPTIDES AND
TITLE OF INVENTION: POLYPEPTIDES
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/173,508

```

: FILING DATE: 23-DEC-1993
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 18740/135/CACO
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202 672 5300
: TELEFAX: 202 672 5399
: TELEX: 904136
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2185 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 531..2069
: FEATURE:
: NAME/KEY: sig_peptide
: LOCATION: 531..902
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: 903..2069
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 531..533
: OTHER INFORMATION: /note="Met at position -124"
: OTHER INFORMATION: represents fmet"
:
: US-08-173-508-3

```

```

Query Match          3.4%; Score 42.4; DB 1; Length 2185;
Best Local Similarity 47.4%; Pred. No. 0.093;
Matches 127; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

```

```

QY 524 CCGCGGTGCTGATGGCGGGGCGGCGGTGCGAACCGGCCGACGTCGTGATCGG 583
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1250 CGCGGTGCGGATGACCGCGGTGGCCACCGGGGTGAAGGTGCAAGGTCTCCAA 1309
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 584 CGCGGGACCGCGCGGTGACGACGACCGCGGATCGCCCAAGCGGCGGCGGATTC 643
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1310 CCGCGACGGCTTCTTCTACGACGACGACGCGGTGCTGCGGCTTCAATGTCGGCGGACGA 1369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 644 GGTCTAGACATCAACATGCAAACTTCGCAACTGACCGCGGAGTTCTCGCGCGGAT 703
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1370 CGGCGTCGACGTGACCAACACAGCTATTACACGACCGCGGTGCTTCACTCACTGCAAGGA 1429
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 704 CCACACTGCTACTGATGCGGCGTACGAGCTGAGAGGTCGCGTCAAAAGTGGCGACTGGT 763
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1430 CGACCCGACGAGAAAGGCGCTGTCGAGGCGGCTTCGCGGCGCTCCGGTACGCGGAGAA 1489
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 764 GATTGGGCGCTCTGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 791
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1490 GAAGGCGCGGTCAACGTCGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1517
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 15
US-08-265-310-3
: Sequence 3, Application US/08265310
: Patent No. 5856166
: GENERAL INFORMATION:
: APPLICANT: Bartfield, Daniel
: APPLICANT: Butler, Michael J.
: APPLICANT: Hadary, Dany
: APPLICANT: Jemish, David
: APPLICANT: Krieger, Timothy
: APPLICANT: Malek, Lawrence T.
: APPLICANT: Soostmeyer, Gisela
: APPLICANT: Malczyk, Eya
: APPLICANT: Kraysman, Phyllis
: APPLICANT: Garven, Shelia

```

```

: TITLE OF INVENTION: STREPTOCOCCUS PROTEASES AND IMPROVED
: TITLE OF INVENTION: STREPTOCOCCUS STRAINS FOR EXPRESSION OF PEPTIDES AND
: TITLE OF INVENTION: POLYPEPTIDES
: NUMBER OF SEQUENCES: 21
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 3000 K Street, N.W.
: CITY: Washington, D.C.
: COUNTRY: USA
: ZIP: 20007-5109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/265,310
: FILING DATE: 24-JUN-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/173,508
: FILING DATE: 23-DEC-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 18740/133/CACO
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202 672 5300
: TELEFAX: 202 672 5399
: TELEX: 904136
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2185 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 531..2069
: FEATURE:
: NAME/KEY: sig_peptide
: LOCATION: 531..902
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: 903..2069
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 531..533
: OTHER INFORMATION: /note="Met at position -124"
: OTHER INFORMATION: represents fmet"
:
: US-08-265-310-3

```

```

Query Match          3.4%; Score 42.4; DB 3; Length 2185;
Best Local Similarity 47.4%; Pred. No. 0.093;
Matches 127; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

```

```

QY 524 CCGCGGTGCTGATGGCGGGGCGGCGGTGCGAACCGGCCGACGTCGTGATCGG 583
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1250 CGCGGTGCGGATGACCGCGGTGGCCACCGGGGTGAAGGTGCAAGGTCTCCAA 1309
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 584 CGCGGGACCGCGCGGTGACGACGACCGCGGATCGCCCAAGCGGCGGCGGATTC 643
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1310 CCGCGACGGCTTCTTCTACGACGACGACGCGGTGCTGCGGCTTCAATGTCGGCGGACGA 1369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 644 GGTCTAGACATCAACATGCAAACTTCGCAACTGACCGCGGAGTTCTCGCGCGGAT 703
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1370 CGGCGTCGACGTGACCAACAGCTATTACACGACCGCGGTGCTTCACTCACTGCAAGGA 1429
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 704 CCACACTGCTACTGATGCGGCGTACGAGCTGAGAGGTCGCGTCAAAAGTGGCGACTGGT 763
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1430 CGACCCGACGAGAAAGGCGCTGTCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1489
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Fri Jun 23 09:31:39 2000

us-09-362-485-8.mri

Page 12

QY 764 GATTGGGCGCTGTGCCAGGCC 791
 | | | | | | |
Db 1490 GAAGGCGCGTCAACGTCCGCGGCC 1517

Search completed: June 22, 2000, 15:16:28
Job time: 18466 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 12:07:49 ; Search time 5541.94 Seconds
(without alignments)
903.245 Million cell updates/sec

Title: US-09-362-485-8
Perfect score: 1235
Sequence: 1 ATCTTCAGATGATATCGAAC.....GGGAGCGATGATGCGGCC 1235

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues
Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST: *
1: em_est1: *
2: em_est2: *
3: em_est3: *
4: em_est4: *
5: em_est5: *
6: em_est6: *
7: em_est7: *
8: em_est8: *
9: em_est9: *
10: em_est10: *
11: em_est11: *
12: em_est12: *
13: em_est13: *
14: em_est14: *
15: em_est15: *
16: em_est16: *
17: em_est17: *
18: em_est18: *
19: em_est19: *
20: em_est20: *
21: em_est21: *
22: em_est22: *
23: em_est23: *
24: em_est24: *
25: em_est25: *
26: em_est26: *
27: em_est27: *
28: em_est28: *
29: em_est29: *
30: em_est30: *
31: em_est31: *
32: em_est32: *
33: em_est33: *
34: em_est34: *
35: em_est35: *
36: em_est36: *
37: em_est37: *
38: em_est38: *
39: em_est39: *
40: em_est40: *
41: em_est41: *
42: em_est42: *
43: em_est43: *
44: em_est44: *

45: gb_est26: *
46: gb_est27: *
47: gb_est28: *
48: gb_est29: *
49: gb_est30: *
50: gb_est31: *
51: gb_est32: *
52: gb_est33: *
53: gb_est34: *
54: gb_est35: *
55: gb_est36: *
56: gb_est37: *
57: gb_est38: *
58: gb_est39: *
59: gb_est40: *
60: gb_est41: *
61: gb_est42: *
62: gb_est43: *
63: gb_est44: *
64: gb_est45: *
65: gb_est46: *
66: gb_est47: *
67: gb_est48: *
68: gb_est49: *
69: gb_est50: *
70: gb_est51: *
71: gb_est52: *
72: gb_est53: *
73: gb_est54: *
74: gb_est55: *
75: gb_est56: *
76: gb_est57: *
77: gb_est58: *
78: gb_est59: *
79: gb_est60: *
80: gb_est61: *
81: gb_est62: *
82: gb_est63: *
83: gb_est64: *
84: gb_est65: *
85: gb_est66: *
86: gb_est67: *
87: gb_est68: *
88: gb_est69: *
89: gb_est70: *
90: gb_est71: *
91: gb_est72: *
92: gb_est73: *
93: gb_est74: *
94: gb_est75: *
95: gb_est76: *
96: gb_est77: *
97: gb_est78: *
98: gb_est79: *
99: gb_est80: *
100: gb_est81: *
101: gb_est82: *
102: gb_est83: *
103: gb_est84: *
104: gb_est85: *
105: gb_est86: *
106: gb_est87: *
107: gb_est88: *
108: gb_est89: *
109: gb_est90: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES

No.	Score	Match	Length	DB	ID	Description
1	57.4	4.6	925	82	CNS0091P	AL053013 Drosophila
2	54	4.4	925	82	CNS0091P	AL053013 Drosophila
3	51.8	4.2	925	82	CNS0091P	AL053013 Drosophila
4	49	4.0	910	82	CNS0060N	AL056629 Drosophila
5	46.6	3.8	844	82	CNS0052P	AL056652 Drosophila
6	45.6	3.7	846	82	CNS0108J	AL099337 Drosophila
7	45.4	3.7	591	69	AM128786	AM128786 fe37d05.Y
8	45.4	3.7	935	82	CNS0066K	AL060051 Drosophila
9	45.2	3.7	1101	83	CNS0175Y	AL108446 Drosophila
10	45	3.6	315	51	AI171257	AI171257 605069D09
11	45	3.6	846	82	CNS0108J	AL099337 Drosophila
12	44.8	3.6	1101	83	CNS0181E	AL108764 Drosophila
13	44.8	3.6	932	82	CNS00720	AL066742 Drosophila
14	44.2	3.5	935	82	CNS0066K	AL066051 Drosophila
15	43.8	3.5	427	51	AI173678	AI173678 sb32B04.Y
16	43.8	3.5	512	74	AV396794	AV396794 AV396794
17	43.8	3.5	521	74	AV397116	AV397116 AV397116
18	43.8	3.5	526	74	AV394827	AV394827 AV394827
19	43.8	3.5	744	83	CNS0172K	AL108698 Drosophila
20	43.6	3.5	1101	83	CNS016DE	AL107216 Drosophila
21	43.4	3.5	414	80	AM286084	AM286084 Lc1_261.B
22	43.4	3.5	465	80	AM287007	AM287007 .L1_264.C
23	43.2	3.5	493	91	AO864063	AO864063 ndeb0022M
24	43.2	3.5	798	82	CNS000AUA	AL055851 Drosophila
25	43	3.5	413	74	AV396947	AV396947 AV396947
26	42.6	3.4	645	82	CNS01213	AL101589 Drosophila
27	42.4	3.4	843	74	CNS00C51	AL059666 Drosophila
28	41.8	3.4	530	74	AV396353	AV396353 AV396353
29	41.8	3.4	932	82	CNS00720	AL066742 Drosophila
30	41.6	3.4	506	74	AM180713	AM180713 MGA0863F
31	41.6	3.4	870	82	CNS0006E2	AL064271 Drosophila
32	41.4	3.4	1101	82	CNS00LXJ	AL078875 Drosophila
33	41.2	3.3	639	60	AI812194	AI812194 605086F10
34	41.2	3.3	774	105	AO327331	AO327331 nbxb0040D
35	41.2	3.3	910	82	CNS0060N	AL066629 Drosophila
36	41	3.3	605	46	AI461529	AI461529 46018G05
37	41	3.3	1101	83	CNS0175Y	AL108446 Drosophila
38	40.8	3.3	384	64	AM054773	AM054773 w59d08.x
39	40.8	3.3	385	21	D48746	D48746 R1C51517A
40	40.8	3.3	426	40	C91740	C91740 C91740 Rice
41	40.8	3.3	437	74	AV395790	AV395790 AV395790
42	40.8	3.3	839	82	CNS0049B	AL054280 Drosophila
43	40.8	3.3	964	82	CNS003MG	AL062254 Drosophila
44	40.6	3.3	506	74	AV396405	AV396405 AV396405
45	40.4	3.3	574	91	AO848186	AO848186 LMAJFV1_1

ALIGNMENTS

RESULT 1
LOCUS CNS0091P/c
DEFINITION Drosophila melanogaster genome survey sequence TERT3 end of BAC # BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL053013
VERSION AL053013.1
KEYWORDS GI:4934461
SOURCE fruit fly.
ORGANISM Drosophila melanogaster

REFERENCE 1
AUTHORS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : sequefgenoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a

collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org/TheBDGP/Drosophila_melanogaster_BAC_library.html. The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Osoegawa and Aaron Moser in the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1. 925
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACR19D16"
/note="end : TERT3"
BASE COUNT 120 a 61 c 61 g 172 t 511 others
ORIGIN

Query Match

Best Local Similarity 11.9%; Pred. No. 0.057;
Matches 43; Conservative 176; Mismatches 142; Indels 0; Gaps 0;

QY 441 CCCCTGCTTCCCGATGACGAGTCCGCTGACTCCCGCCAGTTGCGCTTAC 500
DB 920 SCSCSCSBCSSSMTSSSBCSSSBSSTSSSBSSTSSSBSSTSSSBSSTSSS 861
QY 501 CACCTGATCGCAACGAGCGCGCGCTGATGCGCGCGCGCGCGCGCGCGCGA 560
DB 860 SACKKAASSCGCGCGGACBCMCSSSSSCGSASRGRVRSRGGAGRGSGGASA 801
QY 561 CGGCGCAGCTGCTGATGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 620
DB 800 SHSSSSACBSSSSSCASMSASSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 741
QY 621 AAGCGCATGGCG 660
DB 740 VSSASSSSSSSCSSSVSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 681
QY 681 GACCGCAGTCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 740
DB 680 SCRTSASMSAARSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 621
QY 741 GCCGTCAACGCTGCCGACATGCTGATGCGCGCGCGCGCGCGCGCGCGCG 800
DB 620 GGGSVASSSGMSVSSSSGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 561
QY 801 A 801
DB 560 R 560

RESULT 2
LOCUS CNS0091P
DEFINITION Drosophila melanogaster genome survey sequence TERT3 end of BAC # BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL053013
VERSION AL053013.1
KEYWORDS GI:4934461
SOURCE fruit fly.
ORGANISM Drosophila melanogaster

REFERENCE 1
AUTHORS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

COMMENT

- Web : www.genoscope.cns.fr
 Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammose in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1. 925
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone_lib="RPCI-98"
 /clone="BAC19016"
 /note="end : TER3"

BASE COUNT 120 a 61 c 61 g 172 t 511 others
ORIGIN

Query Match 4.4%; Score 54; DB 82; Length 92;

Best Local Similarity 11.8%; Pred. No. 0.26; Mismatches 138; Indels 0; Gaps 0;

Matches 41; Conservative 169; Mismatches 138; Indels 0; Gaps 0;

QY 424 CCGCCGAGCGCAGTACCCCTGCTGCCCCGATGAGCGAAGTCGCGTCACTGCGCC 483

DB 577 SCGCCSSCCBCBCCCCSSSYCCSSBSKSSSTSBSCSCCCSKSCGTSKSSSSS 636

QY 484 CCGAGTTGGCGGTACACCTATGCGAAGCGGCGCGCGTGTCTATGGGCG 543

DB 637 CSSSSSTSTSSSTSSSTSSSSSYTTSKSTASGSGSWAGCGSGTGTST 696

QY 544 GGGTGGCGCGCTGCAACCGCGCGTCTGTGATCGCGCGCGCGGCGGCGTACA 603

DB 697 SSSSSSTSTSSSVSSGSKSTBSGSSSSSSSTSSSBSTSTSSSSSTSSS 756

QY 604 ACGAGCGCCGATGCGCAAGCGGCGCGATGCGTTCATGACATCAATCG 663

DB 757 TCSCCTCCCTSYSSSTSSSTSSSTSSSVTSSSTSSSTSCCCCTMCCTCS 816

QY 664 ACAAACTTCGCACTGCGAGCGAGTTCGCGCGCGATCAACACTGCTACTCATGG 723

DB 817 TYMBEYTSYSCGSSSSSKGVTGCGCGCGSSSTNGMBGTSSACSSSSSSSSVS 876

QY 724 CCAACGAGCTCGAGGTCGCTCAACAGTCCCACTGGTGGTGGG 771

DB 877 SSSKSSSSSSSVSSGSGVSNSSSSKSSSSGSGSVSSGSGSGSVS 924

RESULT 3
LOCUS AV390505 535 bp mRNA
DEFINITION AV390505 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
ACCESSION AV390505
VERSION AV390505.1 GI:6544721
KEYWORDS EST
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE 1 (bases 1 to 535)

AUTHORS Asamizu, E., Nakamura, Y., Sato, S., Fukuzawa, H. and Tabata, S.
TITLE A Large Scale Structural Analysis of cDNAs in a Unicellular Green
JOURNAL Algae, Chlamydomonas reinhardtii. I. Generation of 3433
 Non-redundant Expressed Sequence Tags
 DNA Res. 6, 369-373 (1999)
 On Dec 20, 1995 this sequence version replaced gi:1135919.

COMMENT Contact: Ynakazu Nakamura
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yama 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: ynakazu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.

FEATURES

source

1. 535
 /organism="Chlamydomonas reinhardtii"
 /strain="C9"
 /db_xref="taxon:3055"
 /clone_lib="CM052B05_r"
 /clone="CM052B05_r"
 /dev_stage="photoautotrophic growth"
 /note="vector: p Bluescript II SK-; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 107 a 195 c 148 g 85 t
ORIGIN

Query Match 4.2%; Score 51.8; DB 74; Length 535;

Best Local Similarity 51.5%; Pred. No. 0.63; Mismatches 112; Indels 0; Gaps 0;

Matches 119; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 564 GCCGAGTCGTGATGCGCGCGCGCACCGCGGCTACAGACGACCGCGATCGCAAC 623

DB 221 GCGAGAGAGTGGCCATGACAGACCTCTCTCTCCAGTACGAGCGTGCAGCGGCTG 280

QY 624 GGCATGGCGCGACCGTTCATGACATCAACATGCAAACTTGGCACTGCAC 683

DB 281 CGCTGCGCTGACGCTGCTGCGCAAGCGCTCAAGTACGACGCGGAGCGGCTGAC 340

QY 684 GCGGATTTGGGGCGCGGATGCAACATCGCTACTCATCGGCTTACAGTGGAGTCC 743

DB 341 GCGGAGTGAACCGGCTGATGACACCTCTTCAACGCGGCTTCAACATCGTGGAGCG 400

QY 744 GTCAACGTGCGCGACCTGATTTGGGCGGCTGCTGTCAGCGGCGCAAG 794

DB 401 GCGGCTGTGCGCTGCTGCGCGGAGGCGCTCAAGAGATTAAGCGCATG 451

RESULT 4
LOCUS CNS0060N 910 bp DNA
DEFINITION CNS0060N 910 bp DNA GSS 03-JUN-1999
 Drosophila melanogaster genome survey sequence T7 end of BAC #
 BAC14021 of RPCI-98 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.

ACCESSION AL065629
VERSION AL065629.1 GI:494698
KEYWORDS GSS

SOURCE fruit fly.
ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 910)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

- Web : www.genoscope.cns.fr
 Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammose in Pieter de Jong's laboratory in the Department of

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

Location/Qualifiers
1. .910
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACR14J21"
/note="end : T7"

BASE COUNT 202 a 63 c 112 g 198 t 335 others
ORIGIN

Query Match 4.0% Score 49; DB 62; Length 910;
Best Local Similarity 20.5% Pred. No. 2.4; Mismatches 85; Indels 0; Gaps 0;
Matches 47; Conservative 97; Mismatches 85; Indels 0; Gaps 0;

QY 392 CACCACTGATGCTGCTAGAGACCGCTCAGACGCGGAGGCGCACTACCTGCTTGC 451
Db 713 SRVMSAGSCTASSSSSSASASSASASASASASASASASASASASASASASASAS 654
QY 452 CCGCATGAGCGAATGCGCGGCTGCTGAGTCCGCGGAGTTGCGCTTCCACTGATGG 511
Db 653 VMAVVAASSSSVASGSSSVSSSSRCCASSSASASSSSSSSSSSSSSSSSSSSSSS 594
QY 512 AACCCAGAGGGGCGCGGCTGCTGATGCGGCGGCGGCTGCAACCGCGGAGCT 571
Db 593 CASSSSRSGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 534
QY 572 CCGTGTGATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 620
Db 533 CASCSCGCGCCSSSCSSCCGCCCMCAAMAHACCCSCCCGCCDC 485

RESULT 5

CNS0052P/c 844 bp DNA GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR14J21 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL056652.1 GI:4932342
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 844)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osogawa and
Aaron Mammot in Peter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or

filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

Location/Qualifiers
1. .844
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACR14J21"
/note="end : TET3"

BASE COUNT 261 a 112 c 92 g 35 t 344 others
ORIGIN

Query Match 3.8% Score 46.6; DB 82; Length 844;
Best Local Similarity 15.8% Pred. No. 6.8; Mismatches 146; Indels 0; Gaps 0;
Matches 58; Conservative 146; Mismatches 146; Indels 0; Gaps 0;

QY 886 AAGCTCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 945
Db 733 MMTTCVMAVMAVMAVMAVMAVMAVMAVMAVMAVMAVMAVMAVMAVMAVMAVMA 674
QY 946 ACTGCTGCGCAATGCGCGGCTGCTGAGTCCGCGGAGTTGCGCTTCCACTGATGG 1005
Db 673 THTNSNTTMTTMTTMTTMTTMTTMTTMTTMTTMTTMTTMTTMTTMTTMTTMT 614
QY 1006 CGACATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1065
Db 613 SSSSSSSGSSSBYSBCCBCCBCCBCCBCCBCCBCCBCCBCCBCCBCCBCCBCC 554
QY 1066 ATCCGCGACTACCAAGAGCTTTTTCAGACGACGACGACGACGACGACGACG 1125
Db 553 SSSSBCCSBSSBCCBCCBCCBCCBCCBCCBCCBCCBCCBCCBCCBCCBCC 494
QY 1126 CCACGACCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1185
Db 493 GCKSGKGGKSGKSGKSGKSGKSGKSGKSGKSGKSGKSGKSGKSGKSGKSGKSG 434
QY 1186 CTCGTTACCGGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1235
Db 433 GSSSKTSSSSCBSSGSSSTGSSSKKYSYSTSSSSSCGCTGKNTTTCV 384

RESULT 6

CNS010R/c 846 bp DNA GSS 26-JUL-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC
DEFINITION BACN04N13 of Drosophila library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL099337.1 GI:5619548
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 846)
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
<http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billand at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos of Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

FEATURES

source

Location/Qualifiers
1. .846
/organism="Drosophila melanogaster"

OY	388	CCGGACAGCCTCAATTGCTATCGAGACCCTCCAGACGCCGAAGGGCATTAACCCCTGC	44.7
Db	308	CCCCCAYVVMYGMMAAABCCCACAAASCSGCCSCSSCCSGCCGSCCSMASCSGCSGCSG	24.9
OY	448	TTGCCCCGATGACGCAATGTCGCCGTGCATCTGCGCCGCCAAGTGGGCTTACCACTGA	50.7
Db	248	AAGMGSMSTGVGGACAAGGGCGTAGAACCCSCCACSCCSSSCSACSSCSSSSSS	18.9
OY	508	TGCGAATCCCAAGGGGGCCGCGGTGTGCTGATGGGGCGGGGTGCCGGGGTGGAACCCGCCG	56.7
Db	188	CCTGACCCCCSSAAGGGGCCSCSSCCGCCCAACCCCGCCSGACVCATCGCCAC - CMMVGGGSG	13.0
OY	568	ACGTGCTGTATCGGCGCCGCGCAACCGCGGCTTCAACGCAAGCCCGCATCGCCAAGCA	62.7
Db	129	ACGGAAAAGGAGAAGGGCGCGGNACCCCGCCACACCCCGCCCGCCGACAGCGCGCGGCA	7.0
OY	628	TGGGGCGCGACG	63.8
Db	69	AGCGCCCAAC	5.9

LOCUS	DEFINITION	FEATURES
CNS00720		
CNS00720	932 bp	DNA
CNS00720	Drosophila melanogaster genome survey sequence T7 end of BAC # BR01A09 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	GSS
		03-JUN-1999

ACCESSION	AL066742
VERSION	AL066742.1
	GI:4945205

KEYWORDS	GSS.
SOURCE	fruit fly.

ORGANISM:

REFERENCE

AUTHORS
TITLE
JOURNAL

COMMENT

Submitted (22-JUN-1999) Genoscope - Centre National de Séquençage
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Oosawa and
Aaron Mammoxer in Pieret de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPci-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP for partial
isogenic strain y2; cn bw sp. the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

Source

BASE COUNT	155 a	202 c	241 g	91 t	243 others
ORIGIN					

Query Match	3.68;	Score 44.8;	DB 82;	Length 932;
Best Local Similarity	34.1%;	Pred. No. 15;		
Matches 73;	Conservative 47;	Mismatches 94;	Indels 0;	Gaps 0

[illegible]

135 859

RESULT 14			
CNS006XK/c			
LOCUS	CNS006XK	935 bp	DNA
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC #		
		GSS	03-JUN-1999

BACR14N09 of RPCT-98 library from *Drosophila melanogaster* (fruit fly), genomic survey sequence.

ACCESSION	AL066051
VERSION	AL066051.1
KEYWORDS	GI:4945019
CCC	

ALIMORDS
SOURCE
ORGANISM

REFERENCE
1 (phases 1 to 935)

AUTHORS
TITLE
JOURNAL

COMMENT

Sub.itted: 02-JUN-1999) Genoscope - Centre National de Séquençage
BP 11 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr
Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org/TheBDGP/Drosophila>
melanogaster BAC library was prepared by Kazuhiro Oosawa and
Aaron Mammossier in Pieret de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPc1-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y¹; cn bw sp. the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACpac Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

Source

BASE COUNT	257 a	170 c	162 g	96 t	250 others
ORIGIN					

Query Match	3.68;	Score 44.2;	DB 82;	Length 935;
Best Local Similarity	27.48;	Pred. No. 20;		
Matches 93;	Conservative 83;	Mismatches 163;	Indels 0;	Gaps 0

QY 406

db 856 C C G C G C G S S S G C C C C S G S C G C C C C S G S C S C S S C C C G C C C G S C S C G C G C G C 79

QY 466 TCGCCGTCGACTCGCCCGCCAGGTTGGCGCTTACCACTGATGCAACCAAGGGGCC 525

Db 796 SGGCCSGGCGCGCGGCCGCCGSKKCGCCGGGGSCSSSSGGSGCGGGGG 737

FEATURES	source
LOCUS	Al1736678
DEFINITION	sb52b04.y1 Gm-c1012 glycine max CDNA clone GENOME SYSTEMS CLONE ID Gm-c1012-104 5' similar to TR.O22253 O22253 PUTATIVE PHOTOLYASE/BLUE-LIGHT RECEPTOR., mRNA sequence.
ACCESSION	Al1736678
VERSION	Al1736678.1 GI:5058202
KEYWORDS	EST.
SOURCE	soybean.
ORGANISM	glycine max
REFERENCE	1 (bases 1 to 427) Shoemaker, R., Keim, P., Vodkin, L., Eprelding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Pearson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
TITLE	Public soybean EST project
JOURNAL	unpublished (1999)
COMMENT	On May 18, 1998 this sequence version replaced gi:3138249. Contact: Shoemaker R/Public Soybean EST Project Public Soybean EST project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.edu This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com Seq primer: -40RP from Glbco High quality sequence stop: 427. Location/Qualifiers 1..427

synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XL10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelidng.

Query Match	3.5%;	Score 43.8;	DB 51;	Length 427;
Best Local Similarity	45.3%;	Pred. No. 21;		
Matches 159;	Conservative 0;	Mismatches 192;	Indels 0;	Gaps 0;
QY 581	CGGGCGCGGCACCGCGGCTACAAAGCAGCCCGCATCGCCAAAGCGCATGGCGACCGT	640		
Db 34	CGCCCTCCGGCGCGCGCGCGCTGTGTGGTTGCGCAAGCATCTCGGCTCCTCGCAATGA	93		
QY 641	TACGGTTTACACATCACATCTGCACAACTTCGGCACTCGACGCCGAGTTTGGGGCG	700		
Db 94	GTCGCTCACCCCGCGCAACAAAGACTCCCTCTCCGCTCCCGCTACTCGTTAGACCC	153		
QY 701	GATCCACACCTGCTACTCATCGGCTTAGAGCTGAGGCTGCGTCAACAGCGGACCT	760		
Db 154	CTCCGACTACGGCAGAGTGGCATCCGGCTTGACACAGACCGGCCCTTACCGGGCGCCT	213		
QY 761	GGTGATTGGGGCGGCTCTGATGTCAGGCGGCCAACAGCACCAATTATAGTCTCAATTACT	820		
Db 214	CCTCATCGACCTCGCTCTCGAACCTCCGGCGGACGCTCCAGGGGGGGGTCGATCTGTT	273		
QY 821	TGTGCGCATATGAACAACAGTGGGGTACTGGTGGATATAGCCATCGACAGGCGGCTG	880		
Db 274	CGTGGCGCTCGGGGAAGCCCGAGACGGTCTGGTGAGATCTGCCAAGGCGGCTGGTGCGGA	333		
QY 881	TTTGGAAGGCTCAGACCGACCACTAGACACACCCGAGCTTCCGGCGTCA	931		
Db 334	CGCGGTGTAAGCCACCGCGCAGGTTCTTCACGACAGGAGGGAAGGAGGA	384		

Search completed: June 22, 2000, 12:07:56
Job time: 10390 sec

```

/lab-host/XXL0-Gold"
/Note: "vector: pluscript II XR, Site_1: EcoRI, Site_2:
XhoI. This cDNA library was constructed from mRNA isolated
from the apical shoots of 9 to 10 day old etiolated
seedlings. The shoot tips including any emerged leaves
were harvested for RNA isolation. The cDNA library was
prepared using the Stratagene pluscript II XR cDNA
library construction kit. Complementary DNA was

```



```

RESULT 1
LOCUS A87612 1209 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 9 from Patent W09836089.
ACCESSION A87612
VERSION A87612.1 GI:6736252
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1209)
AUTHORS Flohe, L. and Singh, M.
TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL Patent: WO 9836089-A 20-AUG-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
Location/Qualifiers
1..1209
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 228 a 391 c 373 g 216 t 1 others
ORIGIN

```

```

Query Match 99.9%; Score 1208; DB 5; Length 1209;
Best Local Similarity 100.0%; Pred. No. 1,1e-165;
Matches 1209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 ATCTGCAGATTAATCAACTTTTTCACACTGAAGGTACAGTATCGAAGGGGTATC 60
DB 1 ATCTGCAGATTAATCAACTTTTTCACACTGAAGGTACAGTATCGAAGGGGTATC 60
QY 61 ATGGCGCTCGGTATTCGACGAGACCAAAAACAAAGAAATTCGGGTGCGCATCACCCG 120
DB 61 ATGGCGCTCGGTATTCGACGAGACCAAAAACAAAGAAATTCGGGTGCGCATCACCCG 120
QY 121 GCCGGCTCGGGAAGTACACCGCTGTCGTCAGTGAAGTCTCATCGAGGAGTGGCGGA 180
DB 121 GCCGGCTCGGGAAGTACACCGCTGTCGTCAGTGAAGTCTCATCGAGGAGTGGCGGA 180
QY 121 GCCGGCTCGGGAAGTACACCGCTGTCGTCAGTGAAGTCTCATCGAGGAGTGGCGGA 180
DB 121 GCCGGCTCGGGAAGTACACCGCTGTCGTCAGTGAAGTCTCATCGAGGAGTGGCGGA 180
QY 181 GAGGGCTCGGTATTCACGACGCGGATTTCAAGGCGGCGCAACTGGTGGGACCC 240
DB 181 GAGGGCTCGGTATTCACGACGCGGATTTCAAGGCGGCGCAACTGGTGGGACCC 240
QY 241 GCCGACGAGGTGTGGGCGGAGCGTGTATTTATGCTCAAGGTCAGAAAGCATAGCGCG 300
DB 241 GCCGACGAGGTGTGGGCGGAGCGTGTATTTATGCTCAAGGTCAGAAAGCATAGCGCG 300
QY 301 GAATACGCGCGCTGCGACACGCGGCAATCTTGTCAAGTTCCTTTCATTTGGCGGCTCA 360
DB 301 GAATACGCGCGCTGCGACACGCGGCAATCTTGTCAAGTTCCTTTCATTTGGCGGCTCA 360
QY 361 CGTCTTGCACGATGAGTGTGTGATTCGGGCGACGCAATTTGGCTTACGAGACCGTC 420
DB 361 CGTCTTGCACGATGAGTGTGTGATTCGGGCGACGCAATTTGGCTTACGAGACCGTC 420
QY 421 CAGACCGCGGAGCGGAGTACCCCTGCTGCGCCGATGAGCGGAAGTCCGCGTCACTC 480
DB 421 CAGACCGCGGAGCGGAGTACCCCTGCTGCGCCGATGAGCGGAAGTCCGCGTCACTC 480
QY 481 GCCGCGGAGGTGTGGCGCTTACCACTGATGGAACCCAAAGGGGCGCGGTGTCTGATG 540
DB 481 GCCGCGGAGGTGTGGCGCTTACCACTGATGGAACCCAAAGGGGCGCGGTGTCTGATG 540
QY 541 GCGCGGGGTGCGCGGCTGGAACCGGCGAGCTGTGGTATCGGCGCGGACCGCGCGC 600
DB 541 GCGCGGGGTGCGCGGCTGGAACCGGCGAGCTGTGGTATCGGCGCGGACCGCGCGC 600
QY 601 TACAACGAGCGCGCATCGCCAAAGGATGGCGCGGAGCGGTCGTTACGTTTACATCAAC 660
DB 601 TACAACGAGCGCGCATCGCCAAAGGATGGCGCGGAGCGGTCGTTACGTTTACATCAAC 660
QY 661 ATGCACAAACTTGGGCACTGACGCGGAGTTCGGGCGGATTCACACTCGCTACTCA 720

```

```

DB 661 ATGCACAAACTTGGGCACTGACGCGGAGTTCGGGCGGATTCACACTCGCTACTCA 720
QY 721 TCGGCTTACGAGCTCGAGGAGTCCGTCAAAGCTGCCGACCTGGTATTTGGGCGCTCTG 780
DB 721 TCGGCTTACGAGCTCGAGGAGTCCGTCAAAGCTGCCGACCTGGTATTTGGGCGCTCTG 780
QY 781 GTGCGGAGGCGGCAAGGCAACCAATAGTCTGGAATTCATCTGTGGCGATATGAACCA 840
DB 781 GTGCGGAGGCGGCAAGGCAACCAATAGTCTGGAATTCATCTGTGGCGATATGAACCA 840
QY 841 GTGCGGAGTGTGGTATATAGCATGACACGAGGCGGCTGTTGAAAGCTCACAGCCG 900
DB 841 GTGCGGAGTGTGGTATATAGCATGACACGAGGCGGCTGTTGAAAGCTCACAGCCG 900
QY 901 ACCACCTTACGACACCCGAGCGTTCGCGTGCACAGACCGCTGTTTACTGTGGCGAAG 960
DB 901 ACCACCTTACGACACCCGAGCGTTCGCGTGCACAGACCGCTGTTTACTGTGGCGAAG 960
QY 961 ATGCGCGGCTGCGTGGCGGAAGAGTGAAGTACGCGCTGACCAAGCGGAGTGGCGTAT 1020
DB 961 ATGCGCGGCTGCGTGGCGGAAGAGTGAAGTACGCGCTGACCAAGCGGAGTGGCGTAT 1020
QY 1021 GTGCTGAGAGTTCGCGACCATGCGTGGCGGCGCGCTGCGGTGCAATCCGCACTAGCC 1080
DB 1021 GTGCTGAGAGTTCGCGACCATGCGTGGCGGCGCGCTGCGGTGCAATCCGCACTAGCC 1080
QY 1081 AAAGGTCTTTCGACGACGCAAGGCGGCTTACTGTCCGAACGGGTGGCCACCACTGGGG 1140
DB 1081 AAAGGTCTTTCGACGACGCAAGGCGGCTTACTGTCCGAACGGGTGGCCACCACTGGGG 1140
QY 1141 GTGCGGTTACCGAGCGCGGACGCTGTCGCTGACCTCGGCGGCTGTTACGCGGAG 1200
DB 1141 GTGCGGTTACCGAGCGCGGACGCTGTCGCTGACCTCGGCGGCTGTTACGCGGAG 1200
QY 1201 CACGCTCG 1209
DB 1201 CACGCTCG 1209

```

```

RESULT 2
LOCUS A87613 1236 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 10 from Patent W09836089.
ACCESSION A87613
VERSION A87613.1 GI:6736253
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1236)
AUTHORS Flohe, L. and Singh, M.
TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL Patent: WO 9836089-A 20-AUG-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
Location/Qualifiers
1..1236
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 236 a 395 c 385 g 220 t
ORIGIN

```

```

Query Match 99.9%; Score 1208; DB 5; Length 1236;
Best Local Similarity 99.9%; Pred. No. 1,1e-165;
Matches 1208; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 ATCTGCAGATTAATCAACTTTTTCACACTGAAGGTACAGTATCGAAGGGGTATC 60
DB 1 ATCTGCAGATTAATCAACTTTTTCACACTGAAGGTACAGTATCGAAGGGGTATC 60
QY 61 ATGGCGCTCGGTATTCGACGAGACCAAAAACAAAGAAATTCGGGTGCGCATCACCCG 120

```

```

Db      61 ATGGCGGTGGTATTTCGACCGGACCAAAACAGCAATTCCGGGTGGGCATCACCCG 120
QY      121 GCCGGCGTGGGAACTAAACCCGTCGTGGCATGAGTGTCTATCCAGCAGGTGCCGGA 180
Db      121 GCCGGCGTGGGAACTAAACCCGTCGTGGCATGAGTGTCTATCCAGCAGGTGCCGGA 180
QY      181 GAGGGGTGGGCTATCCGACCGGCGGATTCAGAGGCGGAGGCGGACGACGTGGCGACC 240
Db      181 GAGGGGTGGGCTATCCGACCGGCGGATTCAGAGGCGGAGGCGGACGACGTGGCGACC 240
QY      241 GCGGACAGAGTGTGGCGGACGCGTGAATTTATGCTCAAGGTCAAAAGACGATAGCGGG 300
Db      241 GCGGACAGAGTGTGGCGGACGCGTGAATTTATGCTCAAGGTCAAAAGACGATAGCGGG 300
QY      301 GAATACGCGCCGCTGGGACACGGGCGAGATCTTGTCAAGTCTTTGATTTGGCCGCTA 360
Db      301 GAATACGCGCCGCTGGGACACGGGCGAGATCTTGTCAAGTCTTTGATTTGGCCGCTA 360
QY      361 CGTGTTCGACCGATCGTTGTGATTCGCGGACGACGATTCGCTTACGATTTGGCGGCTC 420
Db      361 CGTGTTCGACCGATCGTTGTGATTCGCGGACGACGATTCGCTTACGATTTGGCGGCTC 420
QY      421 CAGACCGCGGACGCGGCGCTACCCCTGCTTGCCTGCGGATG  GCGAAGTCGCGGCTCAGTC 480
Db      421 CAGACCGCGGACGCGGCGCTACCCCTGCTTGCCTGCGGATG  GCGAAGTCGCGGCTCAGTC 480
QY      481 GCGGCGGAGTTGGCGCTTACCACTGATGCGCAACCCAGGGGGCGCGGTGTGCTGATG 540
Db      481 GCGGCGGAGTTGGCGCTTACCACTGATGCGCAACCCAGGGGGCGCGGTGTGCTGATG 540
QY      541 GCGGCGGAGTTGGCGCTTACCACTGATGCGCAACCCAGGGGGCGCGGTGTGCTGATG 600
Db      541 GCGGCGGAGTTGGCGCTTACCACTGATGCGCAACCCAGGGGGCGCGGTGTGCTGATG 600
QY      601 TACACGACGCGCGCATCGGCAAGCGCATGCGCGGACGCTTACGCTTACGATCAAC 660
Db      601 TACACGACGCGCGCATCGGCAAGCGCATGCGCGGACGCTTACGCTTACGATCAAC 660
QY      661 ATGCAAACTTGGGAACTCGAGCGCGGAGTTGTGGCGCGGATCCACATCTGCTACTCA 720
Db      661 ATGCAAACTTGGGAACTCGAGCGCGGAGTTGTGGCGCGGATCCACATCTGCTACTCA 720
QY      721 TCGGCTTACGAGTTCGAGGGTGGCGTCAAAAGTGCAGACCTGGTGAATTGGGGCGTCTCA 780
Db      721 TCGGCTTACGAGTTCGAGGGTGGCGTCAAAAGTGCAGACCTGGTGAATTGGGGCGTCTCA 780
QY      781 GTGCGAGGCGCGCAAGGCGACCAAAATAGTCTGCAATTCACCTTGTGCGCATATGAACA 840
Db      781 GTGCGAGGCGCGCAAGGCGACCAAAATAGTCTGCAATTCACCTTGTGCGCATATGAACA 840
QY      841 GGTGGGCTACTGGTGAATATAGCCATCGACAGGGGGGCGTTCGAAGGCTACAGACG 900
Db      841 GGTGGGCTACTGGTGAATATAGCCATCGACAGGGGGGCGTTCGAAGGCTACAGACG 900
QY      901 ACCACTAGACACCGGACGCTGCGCGTGCAGACAGCGTGTATTTACTGCTGGGGAAC 960
Db      901 ACCACTAGACACCGGACGCTGCGCGTGCAGACAGCGTGTATTTACTGCTGGGGAAC 960
QY      961 ATGCGCGCTCGGTGCGGAGAGAGTGCAGCTGACGCGTGAACAACGCGAGATGCCGTAT 1020
Db      961 ATGCGCGCTCGGTGCGGAGAGAGTGCAGCTGACGCGTGAACAACGCGAGATGCCGTAT 1020
QY      1021 GTGCTGAGGCTTTCGACCAATGCTGCTGGCGGGGCGGTGCGGAGTTCGGAATCGGAGTAC 1080
Db      1021 GTGCTGAGGCTTTCGACCAATGCTGCTGGCGGGGCGGTGCGGAGTTCGGAATCGGAGTAC 1080
QY      1081 AAGGCTTTTCGACGACGAGAGGGGGGTTACTGTCGAAGCGGTGGCCACGACCTGGGG 1140
Db      1081 AAGGCTTTTCGACGACGAGAGGGGGGTTACTGTCGAAGCGGTGGCCACGACCTGGGG 1140
QY      1141 GTGCGGTTACGAGCGCGCGGCGAGCGTGTGCTGACTCTCGGCGCTGTTACGCCGAG 1200
Db      1141 GTGCGGTTACGAGCGCGCGGCGAGCGTGTGCTGACTCTCGGCGCTGTTACGCCGAG 1200

```

```

QY      1201 CACAGCTCG 1209
Db      1201 CACAGCTCG 1209

RESULT 3
A89753
LOCUS      A89753          1236 bp      DNA
DEFINITION Sequence 10 from Patent WO90,2862.
ACCESSION A89753
VERSION    A89753.1 GI:6738287
KEYWORDS
SOURCE
ORGANISM   unidentified.
            unidentified.
REFERENCE  1 (bases 1 to 1236)
AUTHORS    Flohe, L. and Singh, M.
TITLE      L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
JOURNAL    Patent: WO 9832862-A 30-JUL-1998;
            FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
            Location/Qualifiers
            source
                /organism="unidentified"
                /db_xref="taxon:32644"
BASE COUNT 236 a 395 c 385 g 220 t
ORIGIN
Query Match          99.9% Score 1208; DB 5; Length 1236;
Best Local Similarity 99.9%; Pred. No. 1, Le-165;
Matches 1208; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 ATCTTGACAGTTAATGAACTTCTTACACTGAAAGCGTACGATTCGAGAGGGTAATC 60
Db      1 ATCTTGACAGTTAATGAACTTCTTACACTGAAAGCGTACGATTCGAGAGGGTAATC 60
QY      61 ATGCGGTGGGTATTCGCGACCGAGACCAAAACAGAAATTCGGGTGGCATCACCCG 120
Db      61 ATGCGGTGGGTATTCGCGACCGAGACCAAAACAGAAATTCGGGTGGCATCACCCG 120
QY      121 GCGGCGTGGCGGAACTA..CCGTCGTGGCATGAGGTGCTCATCCAGGCGAGTGC CGGA 180
Db      121 GCGGCGTGGCGGAACTA..CCGTCGTGGCATGAGGTGCTCATCCAGGCGAGTGC CGGA 180
QY      181 GAGGGCTCGGCTATACCGACCGGAGTTTCAAGGGCGGACGCGCACTGTGCGGACCC 240
Db      181 GAGGGCTCGGCTATACCGACCGGAGTTTCAAGGGCGGACGCGCACTGTGCGGACCC 240
QY      241 GCGGACAGAGTGTGGCGGACGCTGATTTATTGCTCAAGGTCAAAAGACGATAGCGCG 300
Db      241 GCGGACAGAGTGTGGCGGACGCTGATTTATTGCTCAAGGTCAAAAGACGATAGCGCG 300
QY      301 GAATACGCGCGCTGGGACGCGGAGATCTGTCACGTTCTTGATTTGGCGCGGTCA 360
Db      301 GAATACGCGCGCTGGGACGCGGAGATCTGTCACGTTCTTGATTTGGCGCGGTCA 360
QY      361 GAATACGCGCGCTGGGACGCGGAGATCTGTCACGTTCTTGATTTGGCGCGGTCA 360
Db      361 GAATACGCGCGCTGGGACGCGGAGATCTGTCACGTTCTTGATTTGGCGCGGTCA 360
QY      421 CAGACCGCGGAGGCGGACATACCCCTGCTTGCCTGCGGATAGGGAAGTCCCGGTGAC 480
Db      421 CAGACCGCGGAGGCGGACATACCCCTGCTTGCCTGCGGATAGGGAAGTCCCGGTGAC 480
QY      481 GCGGCGGAGTTGGCGCTTACCACTGATGCGAACCCAGAGGGGGCGCGGTGCTGATG 540
Db      481 GCGGCGGAGTTGGCGCTTACCACTGATGCGAACCCAGAGGGGGCGCGGTGCTGATG 540
QY      541 GCGGCGGAGTTGGCGCTTACCACTGATGCGAACCCAGAGGGGGCGCGGTGCTGATG 600
Db      541 GCGGCGGAGTTGGCGCTTACCACTGATGCGAACCCAGAGGGGGCGCGGTGCTGATG 600

```

QY 601 TACAAAGCAGCCCGATGCGCAAGGCGATGGGCGGACCGCTTACGGTTACATCAAC 660
 Db 601 TACAAAGCAGCCCGATGCGCAAGGCGATGGGCGGACCGCTTACGGTTACATCAAC 660
 QY 661 ATGCAACAACTTCGCGAATCGACGCGGATTCGCGCGGATCCACACTGCTACTCA 720
 Db 661 ATGCAACAACTTCGCGAATCGACGCGGATTCGCGCGGATCCACACTGCTACTCA 720
 QY 721 TCGGCTTACGAGCTGAGGGTGGCTGCTAAACGTCGCGACCTGGTATGGGCGCTCTG 780
 Db 721 TCGGCTTACGAGCTGAGGGTGGCTGCTAAACGTCGCGACCTGGTATGGGCGCTCTG 780
 QY 781 GTGCGAGGCGCCAAAGGCGCAATTAATGCTGAATTCATCTGTGCGGATTAAGAACCA 840
 Db 781 GTGCGAGGCGCCAAAGGCGCAATTAATGCTGAATTCATCTGTGCGGATTAAGAACCA 840
 QY 841 GGTGCGTACTGGTATATAGCATCGACGAGGCGGCTGTTTCGAAAGGCTCAGACCG 900
 Db 841 GGTGCGTACTGGTATATAGCATCGACGAGGCGGCTGTTTCGAAAGGCTCAGACCG 900
 QY 901 ACCACTTACGACACCGGACGCTTCGCGTGCACGACAGCCTGTTTACTGCTGGGAC 960
 Db 901 ACCACTTACGACACCGGACGCTTCGCGTGCACGACAGCCTGTTTACTGCTGGGAC 960
 QY 961 ATGCGCGCTCGGTGCGCAAGAGCTGACCTAGCGCTGACCAAGCGAGATGCGGTAT 1020
 Db 961 ATGCGCGCTCGGTGCGCAAGAGCTGACCTAGCGCTGACCAAGCGAGATGCGGTAT 1020
 QY 1021 GTGCTGAGCTTGGCGACCATGCTGGCGGGCGGCTGCCGCTGCAATCGGCACTAGCC 1080
 Db 1021 GTGCTGAGCTTGGCGACCATGCTGGCGGGCGGCTGCCGCTGCAATCGGCACTAGCC 1080
 QY 1081 AAGGCTTTTCAGCGACGAGAGGCGGCTTACTGTCCGAAGGGGTGCGACCGACTGGG 1140
 Db 1081 AAGGCTTTTCAGCGACGAGAGGCGGCTTACTGTCCGAAGGGGTGCGACCGACTGGG 1140
 QY 1141 GTGCGCTTACCGAGCGCGCGACGCTGCTGGCTGACTCTCGGCGCTGTTACGCGGAG 1200
 Db 1141 GTGCGCTTACCGAGCGCGCGACGCTGCTGGCTGACTCTCGGCGCTGTTACGCGGAG 1200
 QY 1201 CMCAGCTCG 1209
 Db 1201 CMCAGCTCG 1209
 RESULT 4
 A87606
 LOCUS A87606 1235 bp DNA PAT 22-JAN-2000
 DEFINITION Sequence 3 from Patent WO9836089.
 ACCESSION A87606
 VERSION A87606.1 GI:6736246
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1235)
 FLOHE, L. and Singh, M.
 TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
 JOURNAL Patent: WO 9836089-A 20-AUG-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source 1..1235
 Location/Qualifiers
 BASE COUNT 236 a 394 c 382 g 220 t 3 other:
 ORIGIN
 Query Match 99.8%; Score 1207; DB 5; Length 1235;
 Best Local Similarity 99.8%; Pred. No. 1 5e-155;
 Matches 1207; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ATCTTGAGATTATGAACTTCTTCACACTGAGCGTACAGCTATCGAGAGGGTAAATC 60

Db 1 ATCTTGAGATTATGAACTTCTTCACACTGAGCGTACAGCTATCGAGAGGGTAAATC 60
 QY 61 ATGCGCGCTGGTATTCGACCGAGACCAAAACAAAGATTCGCGGTGGCATACCCCG 120
 Db 61 ATGCGCGCTGGTATTCGACCGAGACCAAAACAAAGATTCGCGGTGGCATACCCCG 120
 QY 121 GCCGCGTCCGGGAATTAACCCGTCGTGGCATAAGTGTCTATCCAGGCAAGTGGCGGA 180
 Db 121 GCCGCGTCCGGGAATTAACCCGTCGTGGCATAAGTGTCTATCCAGGCAAGTGGCGGA 180
 QY 181 GAGGCGTGGCTATACCGGAGCGGATTTCAAGGCGGACCGGCAACTGGTGGCAC 240
 Db 181 GAGGCGTGGCTATACCGGAGCGGATTTCAAGGCGGACCGGCAACTGGTGGCAC 240
 QY 241 GCCGACAGCTGTGGGCGGACGCTGATTTATGCTCAAGGTCAAAGAACGATAGCGCG 300
 Db 241 GCCGACAGCTGTGGGCGGACGCTGATTTATGCTCAAGGTCAAAGAACGATAGCGCG 300
 QY 301 GAATAGGCGCGCTGCGACGAGGCGAGATCTTTCAGCTTCTGCAATTTGGCGCGCA 360
 Db 301 GAATAGGCGCGCTGCGACGAGGCGAGATCTTTCAGCTTCTGCAATTTGGCGCGCA 360
 QY 361 GGTGCTTACGCGATGCGTGTGATTCGCGACACGACGTCATTTGCTTACGAGACCGTC 420
 Db 361 GGTGCTTACGCGATGCGTGTGATTCGCGACACGACGTCATTTGCTTACGAGACCGTC 420
 QY 421 CAGACCGCGGACGAGGCGACTTCCCTGCTGCTGCGCGATGACGAGAGTGGCGGTGACTC 480
 Db 421 CAGACCGCGGACGAGGCGACTTCCCTGCTGCTGCGCGATGACGAGAGTGGCGGTGACTC 480
 QY 481 GCCGCGCGGTTGGCGCTTACCATCTGATGCGAACCCAAAGGCGGCGCGGTGCTGATG 540
 Db 481 GCCGCGCGGTTGGCGCTTACCATCTGATGCGAACCCAAAGGCGGCGCGGTGCTGATG 540
 QY 541 GCGGCGGTGGCGCGGCTGGAACCGGCGGACGCTGCTGATGCGGCGGCGGACCGCGCG 600
 Db 541 GCGGCGGTGGCGCGGCTGGAACCGGCGGACGCTGCTGATGCGGCGGCGGACCGCGCG 600
 QY 601 TACAAAGCAGCCCGATGCGCAAGGCGATGGGCGGACCGCTTACGCTTACATCAAC 660
 Db 601 TACAAAGCAGCCCGATGCGCAAGGCGATGGGCGGACCGCTTACGCTTACATCAAC 660
 QY 661 ATGCAACAACTTCGCGAATCGACGCGGATTCGCGCGGATCCACACTGCTACTCA 720
 Db 661 ATGCAACAACTTCGCGAATCGACGCGGATTCGCGCGGATCCACACTGCTACTCA 720
 QY 721 TCGGCTTACGAGCTGAGGGTGGCTGCTAAACGTCGCGACCTGGTATGGGCGCTCTG 780
 Db 721 TCGGCTTACGAGCTGAGGGTGGCTGCTAAACGTCGCGACCTGGTATGGGCGCTCTG 780
 QY 781 GTGCGAGGCGCCAAAGGCGCAATTAATGCTGAATTCATCTGTGCGGATTAAGAACCA 840
 Db 781 GTGCGAGGCGCCAAAGGCGCAATTAATGCTGAATTCATCTGTGCGGATTAAGAACCA 840
 QY 841 GGTGCGTACTGGTATATAGCATCGACGAGGCGGCTGTTTCGAAAGGCTCAGACCG 900
 Db 841 GGTGCGTACTGGTATATAGCATCGACGAGGCGGCTGTTTCGAAAGGCTCAGACCG 900
 QY 901 ACCACTTACGACACCGGACGCTTCGCGTGCACGACAGCCTGTTTACTGCTGGGAC 960
 Db 901 ACCACTTACGACACCGGACGCTTCGCGTGCACGACAGCCTGTTTACTGCTGGGAC 960
 QY 961 ATGCGCGCTCGGTGCGCAAGAGCTGACCTAGCGCTGACCAAGCGAGATGCGGTAT 1020
 Db 961 ATGCGCGCTCGGTGCGCAAGAGCTGACCTAGCGCTGACCAAGCGAGATGCGGTAT 1020
 QY 1021 GTGCTGAGCTTGGCGACCATGCTGGCGGGCGGCTGCCGCTGCAATCGGCACTAGCC 1080
 Db 1021 GTGCTGAGCTTGGCGACCATGCTGGCGGGCGGCTGCCGCTGCAATCGGCACTAGCC 1080
 QY 1081 AAGGCTTTTCAGCGACGAGAGGCGGCTTACTGTCCGAAGGGGTGCGACCGACTGGG 1140
 Db 1081 AAGGCTTTTCAGCGACGAGAGGCGGCTTACTGTCCGAAGGGGTGCGACCGACTGGG 1140

Db 1081 AAGGCTTTGACGACGAGAGGGGCTTACTGTCCGAAAGGGTGGCCAGCCGACCTGGGG 1140
 QY 1141 GGGCGGTTACAGGAGCCGCGGCTGTGGCTGACTGTGCGCGCTGTTACCGGAG 1200
 Db 1141 GTGCGGTTACAGGAGCCGCGGCTGTGGCTGACTGTGCGCGCTGTTACCGGAG 1200
 QY 1201 CMCAGCTCG 1209
 Db 1201 CACACNTCG 1209

RESULT 5
 A89746 1235 bp DNA PAT 22-JAN-2000
 LOCUS A89746 Sequence 3 from Patent WO9832862.
 DEFINITION A89746
 VERSION A89746.1 GI:6738280
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1235)
 FLOHE, L. and Singh, M.
 TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
 JOURNAL Patent: WO 9832862-A 30-JUL-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source 1..1235
 /organism="unidentified"
 /db_xref="taxon:32644"

BASE COUNT 236 a 394 c 382 g 220 t 3 others
 ORIGIN

Query Match 99.8%; Score 1207; DB 5; Length 1235;
 Best local similarity 99.8%; Pred. No. 1.5e-165;
 Matches 1207; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCTTCAGATTATATGAACTTCTTACACTGAAAGCTACAGTATCGAGGGGTAATC 60
 Db 1 ATCTTCAGATTATATGAACTTCTTACACTGAAAGCTACAGTATCGAGGGGTAATC 60
 QY 61 ATGCGGTGCTGATTCGACGACGAGCAAAACAAAGCAATTCGGGGTGGCCATCACCCG 120
 Db 61 ATGCGGTGCTGATTCGACGACGAGCAAAACAAAGCAATTCGGGGTGGCCATCACCCG 120
 QY 121 GCGGCGTGGGGAAGTAACCCGCTGCGCATGAGGTGCTATCCAGCAGGTGCCGA 180
 Db 121 GCGGCGTGGGGAAGTAACCCGCTGCGCATGAGGTGCTATCCAGCAGGTGCCGA 180
 QY 181 GAGGCTGGGCTATACCGGAGCGGATTCGAGGGGCAAGCGCGCACTGGTCCGACC 240
 Db 181 GAGGCTGGGCTATACCGGAGCGGATTCGAGGGGCAAGCGCGCACTGGTCCGACC 240
 QY 241 GCGGACGAGGTGTGGGCGGACGCTGATTTATGCTCAAGGTCAAAAGACGATAGGGCG 300
 Db 241 GCGGACGAGGTGTGGGCGGACGCTGATTTATGCTCAAGGTCAAAAGACGATAGGGCG 300
 QY 301 GAATAGGCGCGCTGCGACACGGGCAAGATTTGTTCAGCTTCTTGATTTGGCGCGTCA 360
 Db 301 GAATAGGCGCGCTGCGACACGGGCAAGATTTGTTCAGCTTCTTGATTTGGCGCGTCA 360
 QY 361 CGAGCTGACGAGGTGTGGTGGATTCGCGGACGACGATTCGATTCGAGACCGTGC 420
 Db 361 CGAGCTGACGAGGTGTGGTGGATTCGCGGACGACGATTCGATTCGAGACCGTGC 420
 QY 421 CAGAGCGCGGAGCGGCTACCCCTGCTTGGCCGATGAGCAAGTCCGCGGTGACTC 480
 Db 421 CAGAGCGCGGAGCGGCTACCCCTGCTTGGCCGATGAGCAAGTCCGCGGTGACTC 480
 QY 481 GCGGCGGAGGTGGCGGTTACCACTGATGGAACCCAAAGGGGCGCGGTGCTGATG 540
 Db 481 GCGGCGGAGGTGGCGGTTACCACTGATGGAACCCAAAGGGGCGCGGTGCTGATG 540

QY 541 GCGGCGGTGCGCGGCTGCGAACCAGCGCGGCTGTGTGATTCGGCGCGGACCGCGGC 600
 Db 541 GCGGCGGTGCGCGGCTGCGAACCAGCGCGGCTGTGTGATTCGGCGCGGACCGCGGC 600
 QY 601 TACAAAGCGAGCGGCTATCGCAAGGCGATGGGCGGACCGGTACGTTTACAGATCAAC 660
 Db 601 TACAAAGCGAGCGGCTATCGCAAGGCGATGGGCGGACCGGTACGTTTACAGATCAAC 660
 QY 661 A1CGCAAACTTCGCGCAACTGACGCGGAGTTCGCGCGGATCCAGCTGCTACTCA 720
 Db 661 ATCGAACAACCTTCGCGCAACTGACGCGGAGTTCGCGCGGATCCAGCTGCTACTCA 720
 QY 721 TCGGCTTACGAGCTTCGAGGCTGCGCTCAAGCTCCGACCTGTTGATTTGGGCGCTCTG 780
 Db 721 TCGGCTTACGAGCTTCGAGGCTGCGCTCAAGCTCCGACCTGTTGATTTGGGCGCTCTG 780
 QY 781 GTGCGAGCGCGCAAGGCGCCCAATTAGTCTGAATTACTTGTGCGCATATGAACA 840
 Db 781 GTGCGAGCGCGCAAGGCGCCCAATTAGTCTGAATTACTTGTGCGCATATGAACA 840
 QY 841 GGTGCGTACTGTGTGATATAGCCATGACAGGCGGCTGTTGGAAGCTCAGACCG 900
 Db 841 GGTGCGTACTGTGTGATATAGCCATGACAGGCGGCTGTTGGAAGCTCAGACCG 900
 QY 901 ACCACTACGACACCCGAGCTTGGCGGTGCGACAGACCGGTTTACTGCTGGTGGAC 960
 Db 901 ACCACTACGACACCCGAGCTTGGCGGTGCGACAGACCGGTTTACTGCTGGTGGAC 960
 QY 961 ATGCCCGCTGCTGCGCAAGAGCTGACCTACGCGCTGACCAACGCGACGATCCGTAT 1020
 Db 961 ATGCCCGCTGCTGCGCAAGAGCTGACCTACGCGCTGACCAACGCGACGATCCGTAT 1020
 QY 1021 GTGCTCGAGCTTGGCGACATGCTGTGGCGGCGGCTGCGGATCCGCGACTAGCC 1080
 Db 1021 GTGCTCGAGCTTGGCGACATGCTGTGGCGGCGGCTGCGGATCCGCGACTAGCC 1080
 QY 1081 AAGGCTTTTGCAGCAGCAAGAGGCGTACTGTCCGAAGGGTGGCCAGCTGAGGG 1140
 Db 1081 AAGGCTTTTGCAGCAGCAAGAGGCGTACTGTCCGAAGGGTGGCCAGCTGAGGG 1140
 QY 1141 GTGCGTTACCGAGCCGCGCAAGCTGTGGCGCTGACTGTGCGCGCTGTTACCGGAG 1200
 Db 1141 GTGCGTTACCGAGCCGCGCAAGCTGTGGCGCTGACTGTGCGCGCTGTTACCGGAG 1200
 QY 1201 CMCAGCTCG 1209
 Db 1201 CACACNTCG 1209

RESULT 6
 A87608 1228 bp DNA PAT 22-JAN-2000
 LOCUS A87608 Sequence 5 from Patent WO9836089.
 DEFINITION A87608
 VERSION A87608.1 GI:6736248
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1228)
 FLOHE, L. and Singh, M.
 TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
 JOURNAL Patent: WO 9836089-A 20-AUG-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source 1..1228
 /organism="unidentified"
 /db_xref="taxon:32644"

BASE COUNT 236 a 391 c 382 g 219 t
 ORIGIN

Db 421 CAGACCGCCGAGGCGCACTACCCCTGCTTGCCTCCGATGAGCGAAGTCCGCGTCACTC 480
 QY 481 GCGGCCAGAGTTGGGCTTACCACTGATGCGCAACCAAGGGGCGCGGTGCTGATG 540
 Db 481 GCGGCCAGAGTTGGGCTTACCACTGATGCGCAACCAAGGGGCGCGGTGCTGATG 540
 QY 541 GCGGGGCTCCGCGCTCGCAACCGGCGAGTGTGTGATGCGGCGCGCAACCGCGG 600
 Db 541 GCGGGGCTCCGCGCTCGCAACCGGCGAGTGTGTGATGCGGCGCGCAACCGCGG 600
 QY 601 TACAAACGACCGCCGATCCCAACGCGCATGGGCGCGACCGTTACGGTTTACATCAAC 660
 Db 601 TACAAACGACCGCCGATCCCAACGCGCATGGGCGCGACCGTTACGGTTTACATCAAC 660
 QY 661 ATCGACAACTTCGGCACTCGACGCGGATTTGGGCGCGATCCACACTGCTACTCA 720
 Db 661 ATCGACAACTTCGGCACTCGACGCGGATTTGGGCGCGATCCACACTGCTACTCA 720
 QY 721 TCGGCTACGAGCTGAGGCTGCGTCAAACTGCGCGACCTGATGATGGGCGCTCTG 780
 Db 721 TCGGCTACGAGCTGAGGCTGCGTCAAACTGCGCGACCTGATGATGGGCGCTCTG 780
 QY 781 GTGCGAGGCGCCAAAGGCGCAAAATTAATGCTCGAATTCATTTGCGCATATGAACA 840
 Db 781 GTGCGAGGCGCCAAAGGCGCAAAATTAATGCTCGAATTCATTTGCGCATATGAACA 840
 QY 841 GGTGCGGTACTGGTGGATATAGCATCGACAGGCGGCTGTTTCAAGGCTTACAGACG 900
 Db 841 GGTGCGGTACTGGTGGATATAGCATCGACAGGCGGCTGTTTCAAGGCTTACAGACG 900
 QY 901 ACCACCTAGACACCAAGGAGTTCGCGCTGCGCGTCAAGACGCTGTTTACTGCGGGAAC 960
 Db 901 ACCACCTAGACACCAAGGAGTTCGCGCTGCGCGTCAAGACGCTGTTTACTGCGGGAAC 960
 QY 961 ATGCGCGCTCGGTGCGGAGAGACGCTGACCTGACGCTGCAACGCGAGATGCGGTAT 1020
 Db 961 ATGCGCGCTCGGTGCGGAGAGACGCTGACCTGACGCTGCAACGCGAGATGCGGTAT 1020
 QY 1021 GTGCTGAGCTTCCGACCATGGCTGGGCGGCGGTGCGCGTCAATCCGCGACTAGCC 1080
 Db 1021 GTGCTGAGCTTCCGACCATGGCTGGGCGGCGGTGCGCGTCAATCCGCGACTAGCC 1080
 QY 1081 AAGGCTTTCGACGACGAGAGGAGGCTTACTGTCCGAAGGCGGTGCGCAACGCTGGGG 1140
 Db 1081 AAGGCTTTCGACGACGAGAGGAGGCTTACTGTCCGAAGGCGGTGCGCAACGCTGGGG 1140
 QY 1141 GTGCGCTTACCGAGCGCGCGACGCTGCTGGCTGACTGCGCGCTGTTACGCGGAG 1200
 Db 1141 GTGCGCTTACCGAGCGCGCGACGCTGCTGGCTGACTGCGCGCTGTTACGCGGAG 1200
 QY 1201 CACACGCTG 1209
 Db 1201 CACACGCTG 1209

RESULT 8
 LOCUS A87607 1237 bp DNA
 DEFINITION Sequence 4 from Patent WO9836089.
 ACCESSION A87607
 VERSION A87607.1 GI:6736247
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1237)
 AUTHORS Flohe, L. and Singh, M.
 TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
 JOURNAL Patent: WO 9836089-A 20-Aug-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source 1..1237
 /organism="unidentified"

BASE COUNT 236 a 394 c 386 g 221 t
 ORIGIN
 Query Match 99.8%; Score 1206.4; DB 5; Length 1237;
 Best Local Similarity 99.8%; Pred. No. 1.9e-165;
 Matches 1207; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCTTGACAGATTAAATGAACTTCTTCACTGAAAGCTACGATATGAGAGGGTAATC 60
 Db 1 ATCTTGACAGATTAAATGAACTTCTTCACTGAAAGCTACGATATGAGAGGGTAATC 60
 QY 61 ATGCGGTGCGTATTCGAGCGAGACCAAAACAAAGAAATTCGGGGTCCATCAACCCG 120
 Db 61 ATGCGGTGCGTATTCGAGCGAGACCAAAACAAAGAAATTCGGGGTCCATCAACCCG 120
 QY 121 GCGGCGTCCGCGAACTAACCGTGTGGCCATGAGTGTCTATCCAGGACAGTCCGGA 180
 Db 121 GCGGCGTCCGCGAACTAACCGTGTGGCCATGAGTGTCTATCCAGGACAGTCCGGA 180
 QY 181 GAGGCTCGGCTATACCGAGCGGATTTCAAGCGCGGCGCAACTGGTCCGACAC 240
 Db 181 GAGGCTCGGCTATACCGAGCGGATTTCAAGCGCGGCGCAACTGGTCCGACAC 240
 QY 241 GCGGACAGGTGGGCGGCGAGCTGATTTATGCTCAAGGTCAAAAGAACGATACGCGG 300
 Db 241 GCGGACAGGTGGGCGGCGAGCTGATTTATGCTCAAGGTCAAAAGAACGATACGCGG 300
 QY 301 GAATACGCGCGCTGCGAGACGCGGAGATCTTTCACATTCCTTGCATTTGGCCGCTCA 360
 Db 301 GAATACGCGCGCTGCGAGACGCGGAGATCTTTCACATTCCTTGCATTTGGCCGCTCA 360
 QY 361 CGTGTTCGACAGATGCGTGTGATTCGCGACAGTCAATTGCTTACAGACCGTC 420
 Db 361 CGTGTTCGACAGATGCGTGTGATTCGCGACAGTCAATTGCTTACAGACCGTC 420
 QY 421 CAGACCGCGAGGCGGCACTACCCGCTTGGCGCGGATGAGGAGTCCGCGTGCATC 480
 Db 421 CAGACCGCGAGGCGGCACTACCCGCTTGGCGCGGATGAGGAGTCCGCGTGCATC 480
 QY 481 GCGGCCAGAGTTGGCGCTTACCACTGATGCGAACCAGAGGGGCGCGGTGCTGATG 540
 Db 481 GCGGCCAGAGTTGGCGCTTACCACTGATGCGAACCAGAGGGGCGCGGTGCTGATG 540
 QY 541 GCGGGGCTCCGCGCTCGAACCAGCGCTGCTGATGCGGCGCGCACCGCGGCG 600
 Db 541 GCGGGGCTCCGCGCTCGAACCAGCGCTGCTGATGCGGCGCGCACCGCGGCG 600
 QY 601 TACAAACGACCGCCGATCCCAACGCGCATGGGCGCGACCGTTACGGTTTACATCAAC 660
 Db 601 TACAAACGACCGCCGATCCCAACGCGCATGGGCGCGACCGTTACGGTTTACATCAAC 660
 QY 661 ATCGACAACTTCGGCACTCGACGCGGATTTGGGCGCGATCCACACTGCTACTCA 720
 Db 661 ATCGACAACTTCGGCACTCGACGCGGATTTGGGCGCGATCCACACTGCTACTCA 720
 QY 721 TCGGCTACGAGCTGAGGCTGCGTCAAACTGCGCGACCTGATGATGGGCGCTCTG 780
 Db 721 TCGGCTACGAGCTGAGGCTGCGTCAAACTGCGCGACCTGATGATGGGCGCTCTG 780
 QY 781 GTGCGAGGCGCCAAAGGCGCAAAATTAATGCTCGAATTCATTTGCGCATATGAACA 840
 Db 781 GTGCGAGGCGCCAAAGGCGCAAAATTAATGCTCGAATTCATTTGCGCATATGAACA 840
 QY 841 GGTGCGGTACTGGTGGATATAGCATCGACAGGCGGCTGTTTCAAGGCTTACAGACG 900
 Db 841 GGTGCGGTACTGGTGGATATAGCATCGACAGGCGGCTGTTTCAAGGCTTACAGACG 900
 QY 901 ACCACCTAGACACCAAGGAGTTCGCGCTGCGCGTCAAGACGCTGTTTACTGCGGGAAC 960
 Db 901 ACCACCTAGACACCAAGGAGTTCGCGCTGCGCGTCAAGACGCTGTTTACTGCGGGAAC 960

QY 961 ATGCCCGCTGGTGGCCGAAGAGCTGACCTGACGCTGACCAAGCAGATGCGCTAT 1020
 DB 961 ATGCCCGCTGGTGGCCGAAGAGCTGACCTGACGCTGACCAAGCAGATGCGCTAT 1020
 QY 1021 GTGCTGACCTGTCGACCATGCTGGCGGGCGGCTGCGATTCGCGACTAGCC 1080
 DB 1021 GTGCTGACCTGTCGACCATGCTGGCGGGCGGCTGCGATTCGCGACTAGCC 1080
 QY 1081 AAGGCTTTCAGCAGCAAGAGGGGCTTACTGTCGAAAGGGGTGGCCACGACCTGGGG 1140
 DB 1081 AAGGCTTTCAGCAGCAAGAGGGGCTTACTGTCGAAAGGGGTGGCCACGACCTGGGG 1140
 QY 1141 GTGCGCTTACCGAGAGCCCGCAGCGCTGCTGAGCTGCTGCGCGCTGTTACGCGGAG 1200
 DB 1141 GTGCGCTTACCGAGAGCCCGCAGCGCTGCTGAGCTGCTGCGCGCTGTTACGCGGAG 1200
 QY 1201 CMCAGCTCG 1209
 DB 1201 CACACGCTG 1209

RESULT 9

LOCUS A89747 1237 bp DNA PAT 22-JAN-2000
 DEFINITION Sequence 4 from Patent WO9832862.
 ACCESSION A89747
 VERSION A89747.1 GI:6738281
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.

REFERENCE
 1 (bases 1 to 1237)
 FLOHE, L. and Singh, M.
 L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
 Patent: WO 9832862-A 30-JUL-1998
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source location/Qualifiers
 1..1237
 /organism="unidentified"
 /db_xref="taxon:32844"

BASE COUNT 236 a 394 c 386 g 221 t
 ORIGIN

Query Match 99.8%; Score 1206.4; DB 5; Length 1237;
 Best Local Similarity 99.8%; Pred. No. 1.9e-165;
 Matches 1207; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCTTCAGATTAATGCAACTTCTTACACTGAGCGTACAGTATGAGAGGGTAAATC 60
 DB 1 ATCTTCAGATTAATGCAACTTCTTACACTGAGCGTACAGTATGAGAGGGTAAATC 60
 QY 61 ATGCGGCTGGTATTCGACGACGACCAAAACAAGAAATTCGGGGTGGCCATCAACCCG 120
 DB 61 ATGCGGCTGGTATTCGACGACGACCAAAACAAGAAATTCGGGGTGGCCATCAACCCG 120
 QY 121 GCGGCGTCCGGAACATCAACCGCTGCGCATGAGTGGCTCATCCAGGAGTGGCGGA 180
 DB 121 GCGGCGTCCGGAACATCAACCGCTGCGCATGAGTGGCTCATCCAGGAGTGGCGGA 180
 QY 181 GAGGCTGGCTATACACGACGCGGATTTAAGGCGCAAGCGGCAACTGGTGGCAAC 240
 DB 181 GAGGCTGGCTATACACGACGCGGATTTAAGGCGCAAGCGGCAACTGGTGGCAAC 240
 QY 241 GCGGACGAGTGGGCGGCGGCTGATTTATGCTCAAGGTCAAAAGACGATAGCGGG 300
 DB 241 GCGGACGAGTGGGCGGCGGCTGATTTATGCTCAAGGTCAAAAGACGATAGCGGG 300
 QY 301 GAATACGCGCGCTGCGACACGCGGAGATCTGTACAGTTCTTGCATTTGGCGCGTCA 360
 DB 301 GAATACGCGCGCTGCGACACGCGGAGATCTGTACAGTTCTTGCATTTGGCGCGTCA 360
 QY 361 CGTGTTCACGAGATGCTGTGTGATTCGGGACCAACGTCATTCGCTACGAGACGTC 420

DB 361 CGTCTTGCACCGATGCTGTGTGATTCGCGCACCAACGATTCGCTACGAGACGTC 420
 QY 421 CAGACCGCGGACGCGGCTTACCCCTGCTTCCCGCATGAGCAATCGCGGTGACATC 480
 DB 421 CAGACCGCGGACGCGGCTTACCCCTGCTTCCCGCATGAGCAATCGCGGTGACATC 480
 QY 481 GCGCGGAGTGGCGCTTACCACTGATGCGAACCAGAGGGGCGCGGTGCTGATG 540
 DB 481 GCGCGGAGTGGCGCTTACCACTGATGCGAACCAGAGGGGCGCGGTGCTGATG 540
 QY 541 GCGGCGTGGCGGCGGCTGCAACCGGCGGAGCTGTGTGATTCGCGCGCGGCGG 600
 DB 541 GCGGCGTGGCGGCGGCTGCAACCGGCGGAGCTGTGTGATTCGCGCGCGGCGG 600
 QY 601 TACAAGCAGCGCGCATCCCAACGAGGATGGGCGGCGCTTACGCTTACGATCAAC 660
 DB 601 TACAAGCAGCGCGCATCCCAACGAGGATGGGCGGCGCTTACGCTTACGATCAAC 660
 QY 661 ATGCAAAACTTGGCACTGACGAGCGGAGTTCTGCGCGGATCCACACTGCTACTCA 720
 DB 661 ATGCAAAACTTGGCACTGACGAGCGGAGTTCTGCGCGGATCCACACTGCTACTCA 720
 QY 721 TCGGCTACGAGCTCGAGGGGCGGCTCAACGTCGACCTGATGAGTGGGCGGCTG 780
 DB 721 TCGGCTACGAGCTCGAGGGGCGGCTCAACGTCGACCTGATGAGTGGGCGGCTG 780
 QY 781 GTGCGAGGCGGCAAGGACCAAAATAGTCTGCAATTCAGTTCGCGCATATGAACA 840
 DB 781 GTGCGAGGCGGCAAGGACCAAAATAGTCTGCAATTCAGTTCGCGCATATGAACA 840
 QY 841 GGTGCGTACTGTGTGATATGCGATGACGAGGCGGCTGTTCCAAAGCTACGACG 900
 DB 841 GGTGCGTACTGTGTGATATGCGATGACGAGGCGGCTGTTCCAAAGCTACGACG 900
 QY 901 ACCACTACGACACCGGAGCTTGGCGGCTGACAGACGCTGTTTACTGCTGCGGAAC 960
 DB 901 ACCACTACGACACCGGAGCTTGGCGGCTGACAGACGCTGTTTACTGCTGCGGAAC 960
 QY 961 ATGCCCGCTCGGTGCC AGAGCTGACCTACGCGCTGACC ACAGGAGATGCGGTAT 1020
 DB 961 ATGCCCGCTCGGTGCCAGAGCTGACCTACGCGCTGACCAACGAGAGATGCGGTAT 1020
 QY 1021 GTGCTGAGCTTGGCGACCATGCTGGGCGGCGGCTGCGATTCGGGCACTAGCC 1080
 DB 1021 GTGCTGAGCTTGGCGACCATGCTGGGCGGCGGCTGCGATTCGGGCACTAGCC 1080
 QY 1141 GTGCGCTTACGAGCGGCGGCGGCTGCTGACCTGCGGCGGCTGCTGAGCGGAG 1200
 DB 1141 GTGCGCTTACGAGCGGCGGCGGCTGCTGACCTGCGGCGGCTGCTGAGCGGAG 1200
 QY 1201 CMCAGCTCG 1209
 DB 1201 CACACGCTG 1209

RESULT 10

LOCUS MTN002 56414 bp DNA BCT 17-JUN-1998
 DEFINITION Mycobacterium tuberculosis H37Rv complete genome; segment 122/162.
 ACCESSION AL008967 AL123456
 VERSION AL008967.1 GI:3261491
 KEYWORDS
 SOURCE Mycobacterium tuberculosis.
 ORGANISM Mycobacterium tuberculosis
 Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 Actinomycetales; Corynebacterinae; Mycobacteriaceae;
 Mycobacterium.
 REFERENCE 1 (bases 1 to 56414)

AUTHORS Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eigmeier K., Gas S., Barry III C.E., Tekla F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holtz R., Hovnsby T., Jagsels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulten J.E., Taylor K., Whitehead S. and Barrall B.G.

TITLE Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence

JOURNAL Nature 393 (6685), 537-544 (1998)

REMARK Eristum: [[published erratum appears in Nature 1998 Nov 12;396(6707):1901]]
2 (bases 1 to 36414)
Parkhill, J.

REFERENCE Direct Submission
Submitted (11-JUN-1998) submitted on behalf of the Mycobacterium tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk
On Jun 27, 1998 this sequence version replaced gi:2624256.

COMMENT Notes:
Details of *M. tuberculosis* sequencing at the Sanger Centre are available on the World Wide Web.
(URL: <http://www.sanger.ac.uk/Projec s/M.tuberculosis/>) CDS have been renumbered from the original cosmid submissions but the old gene designations are in brackets after the new gene numbers.
Gene prediction was based on a Hidden Markov Model of TB genes implemented in TParse (Krogh) supplemented with visual inspection of positional base preference in codons, especially where there is an increase in the observed/expected third position G + C.
CAUTION: In some cases we may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, or ttg) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.
Location/Qualifiers
1. 56414
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
/note=">56414
/note="fragment designated v002. Does not represent a physical clone"
1. 233
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
/clone="Y154"
/complement(3. .527)
/gene="recX"
/complement(3. .527)
/gene="recX"
/note="RV2738c, (MTV002.02c), len: 174 aa, recX, similar to eg. RECX_PSEAE P37860 regulatory protein recX from *Pseudomonas aeruginosa* (153 aa), fasta scores: opt: 161 z-score: 257.2 E(): 3.6e-07, 30.7% identity in 137 aa overlap. Overlaps and extends CDS from overlapping cosmid MTCY154.16c"
/codon_start=1
/transl_table=11
/product="recX"
/protein_id="CA15532.1"
/db_xref="GI:2624257"
/translation="MTVSCPSPSTSRERQAPALICRLTARSRTAPRLAGLAKRGY PEDIGRVLDRLAAGLVDDTPAEQWVOSRRANAKSKRALAEIHLAKGVDDVITTT VIGGIDGAGRGAEKLVARLRREVLIIDGDEARVSRVLVAMLARRGYGTLLACEV VIAELAAERERERV"

gene complement(493. .2865)
/gene="recX"
comple int(493. .2865)
/note="RV2737c, (MTV002.02c), len: 790 aa, recX, identical to RECA_MYCRU P26345 recA protein (790 aa). Contains self-splicing protein element (intron) from 2294 to 969 (c) similar to intron II from TR:E332317 (EMBL:Y13030) DNA-directed DNA polymerase (EC 2.7.7.7) from *Thermococcus* sp. (1829 aa), fasta scores: opt: 81 z-score: 235.2 E(): 6e-06, 24.6% identity in 183 aa overlap. Contains PS00017 ATP/GTP-binding site motif A (P-loop), PS00321 recA signature, and PS00881 protein splicing signature. See Davis et al, (1992) Cell 71(2):201-210"
/codon_start=1
/transl_table=11
/product="recA"
/protc_id="CA15533.1"
/db_xref="GI:2624258"
/db_xref="SWISS-PROT:P26345"
/translation="MTQTPDREKALFLAVNAIEKSGKGYMRUGDEARQPISTYIPG STADLVAGIGLGRVYETIEPSSGKTIYALHAANNQACVAFDAEHALDP DYARKLGVDITSLVSDPTGEQALADMLIRSGALDIVDVAALVPALLEGEM GDSHVGQARLMSQALRMTGALNNSGTATFIRQNRKIGVMGSEETIGVALKF YASVMDVRYEITLKDGTNAVGNRTRVYVNRKGLAETRIEPTGTHRIEDVYG RKP1HVVAADKGTILHARPVVSWFDGTRVDYIGRIAGATVATPDHKLVEGYRA AGELRKQDRAVQPRRFGDSDAP1PDHARLITGLDGDGCVGKTPINFTNOR ALIDVTRAAATLGCARHPQGRISIALAHRRGERNGVADLCQAGITGLAMETITRN WFFEDDIADIVGNILFGLPESDQVSWSEQFARLVGTTTSEQLAHIMHLRLEF GSYVDYPTQKRPSTVNGRIQSKRQVEYFISGMNVYAFASVPMGRNALIQ AIPEXTQGRSGSATYLAAMETAVLYLDERVSTQEAAMAGVSGPGRGMKV IASRLRRDRQALADALDKEDLDMALAEELRYVIREVPTPRARFEDLEELHTL VAEQVYVNCSPKQAEFDIYKKGISREGLSDMGVDGLIRKSCAMPTFBEQJG OGKEARNFLENADVDIEIKIKELIGIAVVTDDPSNDGVLPAEYDF"
complement(790. .2112)
/gene="recA"
/note="recA intron"
complement(790. .807)
/gene="recX"
/note="PS00881 Protein splicing signature"
complement(2197. .2223)
/gene="recA"
/note="PS00321 recA signature"
complement(2644. .2667)
/gene="recX"
/note="PS00017 ATP/GTP-binding site motif A (P-loop)"
complement(2873. .2877)
/note="possible RBS upstream of RV2737c"
complement(3247. .3453)
/gene="RV2738c"
complement(3247. .3453)
/gene="RV2738c"
/note="RV2738c, (MTV002.03c), len: 68 aa, Unknown, N-terminus is highly similar to the N-terminus of the upstream ORF MTV002.07c (78.4% identity in 37 aa overlap); also similar to A10209581 ISC48.5 streptomycin coelicolin cosmid 4H8 (64aa) opt: 185 z-score: 283.5 E(): 2.9e-08; 39.7% identity in 63 aa overlap.
/codon_start=1
/transl_table=11
/product="hypothetical protein RV2738c"
/protein_id="CA15534.1"
/db_xref="GI:2624260"
/db_xref="SPTREMBL:O33281"
/translation="MLAGVRLTEFERHAFHGAAYGSSVLLDHYLTGFDGRSAQAQI EDGVEPRDVMRALCADFDVPRDRW"
complement(3464. .4630)
/gene="RV2739c"
complement(3464. .4630)
/gene="RV2739c"
/note="RV2739c, (MTV002.04c), len: 388 aa, Probable transferase, similar to eg. TR:O51560 (EMBL:U28170) *Pseudomonas aeruginosa* -thamnosyl transferase (426 aa),

fasta_scores: opt: 178 z-score: 226.3 E(): 1.9e-05, 25.9% identity in 425 aa overlap. Equivalent to Mycobacterium leprae protein MLCB3_026 (392 aa); fasta scores 99|294723|MLCB3_2 Mycobacterium leprae cosmid B33 opt: 2112 z-score: 2364.5 E(): 0; 80.9% identity in 388 aa overlap

codon_start=1
/transl_table=11
/product="hypothetical protein Rv2739c"
/protein_id="CA15535.1"
/db_xref="GI:2624261"
/db_xref="SPTREMBL:O33282"
/translation="MRVAVVAGPDGSHSPALALCORPRAADPTLTGYEMLEAAR
AAGIDAVELDLATRDLDAGARIHRRAQAVAVNPRLALEPELVSVATACGG
MAELGIPWEINHPILYLPKSLPISGAGAGTGRRLRDMATLALGSRMRG
LROBAVREVEIGPARDDGPLRLIATLPALEVPDPAPVAVVGPVHPPTDYLA
IPAGTGVVAVVAPSTALGTAGLTVLALQSLTSGTTSVSGSLVYSIADLVTPPA
AVAGLSQAEHLTRADLVICGGHGMVAKTLIAGPVPVAVVQDSQWELIARVVOGS
AVLIRPLADALVAANVNEVLSPPREARARAAVAGAADPVRVCHDALALAG"
4674..5123
/gene="Rv2740"
4674..5123
/note="Rv2740, (MTV002.05), len: 149 aa, unknown, similar
to M. leprae hypothetical protein TR:049850 (EMBL:294723)
MLCB3_03 (B223_F3_140) (178 aa), fasta scores: opt: 498
z-score: 538.9 E(): 7.3e-23, 51.6% identity in 161 aa
overlap"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv2740"
/protein_id="CA15536.1"
/db_xref="GI:2624262"
/db_xref="SPTREMBL:O33283"
/translation="MAELTSPETPEAIRAVEAFNLALONEDPDTVAIGDDL
VYENGFSRIRGGRTATLIRMOGRVFEYKIRHIGDGAVALTEPDAIILGPIRV
OPWCVGEVYDDGRITLIRDFDYIDMKGLIRGLVALVPSLKAFL"
5355..6932
/gene="PE_PGRS"
5355..6932
/note="Rv2741, (MTV002.06), len: 525 aa; Member of M.
tuberculosis PE_PGRS subfamily, similar to many eg.

Query Match 99.8%; Score 1206.4; DB 1; Length 56414;
Best Local Similarity 99.8%; Pred. No. 8.1e-166;
Matches 1207; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

gene
CDS
1 ANCTGCAATTAATGCACTTTCTTCACTGAAAGCGTACAGTATCGAGAGGGTAAATC 60
38201 ATCTTGCAATTAATGCACTTTCTTCACTGAAAGCGTACAGTATCGAGAGGGTAAATC 38260
61 ATGCGCGTGGTATTCGACGACGACAAACAAACGAATTCGCGGTGCGATCAACCCG 120
38261 ATGCGCGTGGTATTCGACGACGACAAACAAACGAATTCGCGGTGCGATCAACCCG 38320
121 GCGGCGTGGCGAACTAACCCGTGTGGCATGAGGTCTCATCCAGGCGAGTCCCGGA 180
38321 GCGGCGTGGCGAACTAACCCGTGTGGCATGAGGTCTCATCCAGGCGAGTCCCGGA 38380
181 GAGGGCTGGCTATACGACGACGCGATTCAGGGCGGCGGCGCAATGGTGGCGAAC 240
38381 GAGGGCTGGCTATACGACGACGCGATTCAGGGCGGCGGCGCAATGGTGGCGAAC 38440
241 GCGGACAGGTGGCGGCGACGCTGATTATTGCTCAACJTCGAAGAACCGATAGCGGCG 300
38441 GCGGACAGGTGGCGGCGACGCTGATTATTGCTCAACJTCGAAGAACCGATAGCGGCG 38500
301 GAATAGCGCGCGCTGCGACAGCGGCGAGATCTTGTCAGCTTTTGCAATTTGGCCGCTCA 360
38501 GAATAGCGCGCGCTGCGACAGCGGCGAGATCTTGTCAGCTTTTGCAATTTGGCCGCTCA 38560
361 CGTGCTGACGATGCGTGTGATTCGGGACACAGCTCAATTTGCCATAGACAGCGTGC 420

Db 38561 CGTCTTACACGATGCGTCTTTGGATTCCGGACCAACGATCAATTTGCTACAGAACCTTC 38620
QY 421 CAGACCGCGGACGCGGCACTACCCCTGCTTCCCGATGAGCGAAGTCCGCGGTGACTC 480
Db 38621 CAGACCGCGGACGCGGCACTACCCCTGCTTCCCGATGAGCGAAGTCCGCGGTGACTC 38680
QY 481 GCGCGCCAGGTTGGCGCTTACCACCTGATGGAACCCAAAGGGGCGCGGTGCTGATG 540
Db 38681 GCGCGCCAGGTTGGCGCTTACCACCTGATGGAACCCAAAGGGGCGCGGTGCTGATG 38740
QY 541 GCGCGGCTGGCGGCGTGAACCGCGGAGCTGCTGTGATCGCGCGCGGACCGCGGCG 600
Db 38741 GCGCGGCTGGCGGCGTGAACCGCGGAGCTGCTGTGATCGCGCGCGGACCGCGGCG 38800
QY 601 TACAACGACGCGCGCATCGCCACGCGATGGGCGGACCGCTTACGTTTACATCAAC 660
Db 38801 TACAACGACGCGCGCATCGCCACGCGATGGGCGGACCGCTTACGTTTACATCAAC 38860
QY 661 ATGACAAACTTGGCAACTGACGCGGAGTTCGGCGGCGGATCCACATCGCTACTCA 720
Db 38861 ATGACAAACTTGGCAACTGACGCGGAGTTCGGCGGCGGATCCACATCGCTACTCA 38920
QY 721 TCGGCTACGAGCTCGAGGCTGCCGTCAACGTCGCGACCTGTGATTGGGCGCGTCTG 780
Db 38921 TCGGCTACGAGCTCGAGGCTGCCGTCAACGTCGCGACCTGTGATTGGGCGCGTCTG 38980
QY 781 GTCCGAGCGCGGACGACCCCAATATGCTCGCAATTCATCTTGCGGCATATGAACCA 840
Db 38981 GTCCGAGCGCGGACGACCCCAATATGCTCGCAATTCATCTTGCGGCATATGAACCA 39040
QY 841 GGTGGGTACTGTGATATATAGCATGACGACGAGGCGGCTGTTGCAAGGCTACGACCG 900
Db 39041 GGTGGGTACTGTGATATATAGCATGACGACGAGGCGGCTGTTGCAAGGCTACGACCG 39100
QY 901 ACCACTACGACCAACCGGAGCTGCGGTGACGACGACGCTGTTTACTGCTGCGGCAAC 960
Db 39101 ACCACTACGACCAACCGGAGCTGCGGTGACGACGACGCTGTTTACTGCTGCGGCAAC 39160
QY 961 ATGCGCGGCTGGTGGCGAAGAGCTGACCTACGCGGTGACCAACGCGGAGATGCCAT 1020
Db 39161 ATGCGCGGCTGGTGGCGAAGAGCTGACCTACGCGGTGACCAACGCGGAGATGCCAT 39220
QY 1021 GTGCTGAGCTTGGCGACATGCGTGGCGGCGGCGGTGATCGGCACTAGCC 1080
Db 39221 GTGCTGAGCTTGGCGACATGCGTGGCGGCGGCGGTGATCGGCACTAGCC 39280
QY 1081 AAAGCTCTTGGACGACGAAAGCGCGCTTACTGTCCGAACGGGTGGCCACGACTGGGG 1140
Db 39281 AAAGCTCTTGGACGACGAAAGCGCGCTTACTGTCCGAACGGGTGGCCACGACTGGGG 39340
QY 1141 GTGCGTTCACGAGCGCGGCTAUCGTGCTGGCGCTGACCTCGCGCGCTGCTTACGGCGAG 1200
Db 39341 GTGCGTTCACGAGCGCGGCTAUCGTGCTGGCGCTGACCTCGCGCGCTGCTTACGGCGAG 39400
QY 1201 CNCAAGTGG 1209
Db 39401 CACACGTGG 39409

RESULT 11
A89752
LOCUS A89752 1208 bp DNA
DEFINITION Sequence 9 from Patent W09832862.
ACCESSION A89752
VERSION A89752.1 GI:6738285
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1208)
AUTHORS Flohe, L. and Singh, M.
TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM

JOURNAL Patent: WO 9832862-A 30-JUL-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES Location/Qualifiers
 source 1..1208
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 228 a 391 c 373 g 216 t
 ORIGIN

Query Match 99.8%; Score 1201; DB 5; Length 1208;
 Best Local Similarity 100.0%; Pred. No. 1,1e-164;
 Matches 1201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTTGAGATTATCGAATCTTCTTCACTGAAAGCGTACAGTATCGAGAGGGTAATC 60
 DB 1 ATCTTGAGATTATCGAATCTTCTTCACTGAAAGCGTACAGTATCGAGAGGGTAATC 60

QY 61 ATGGCGGCGGATTCGACGAGACCAAAACGAATTCGGGGTGGCCATCGACCCCG 120
 DB 61 ATGGCGGCGGATTCGACGAGACCAAAACGAATTCGGGGTGGCCATCGACCCCG 120

QY 121 GCCGCGCTCGGGAACCTAACCCGTCGTGGCCATGAGTGTCTCATCGAGCAGTCCGGA 180
 DB 121 GCCGCGCTCGGGAACCTAACCCGTCGTGGCCATGAGTGTCTCATCGAGCAGTCCGGA 180

QY 181 GAGGCGCTCGGCTATCACGACGCGGATTCAGAGCGGCGCAACTGGTCGGCACC 240
 DB 181 GAGGCGCTCGGCTATCACGACGCGGATTCAGAGCGGCGCAACTGGTCGGCACC 240

QY 241 GCCGACGAGGTGGGCGGAGCGGATTTATGCTCAAGGTCAAGAACCGGATACGGCG 300
 DB 241 GCCGACGAGGTGGGCGGAGCGGATTTATGCTCAAGGTCAAGAACCGGATACGGCG 300

QY 301 GAATACGCGCGCTCGACACGCGGAGATCTTGTACGTTCTTCATTTGGCCCGCTCA 360
 DB 301 GAATACGCGCGCTCGACACGCGGAGATCTTGTACGTTCTTCATTTGGCCCGCTCA 360

QY 361 CGTGCTTACCGATGCGTTGTTGATTCGCGCACGCTCAATTGCCATCAGACCGCTC 420
 DB 361 CGTGCTTACCGATGCGTTGTTGATTCGCGCACGCTCAATTGCCATCAGACCGCTC 420

QY 421 CAGACCGCGGAGCGGCGGACGCTGCTGCGCGGATGAGGAGTGGCGGCTGACATC 480
 DB 421 CAGACCGCGGAGCGGCGGACGCTGCTGCGCGGATGAGGAGTGGCGGCTGACATC 480

QY 481 GCCGCGCGGAGTTGGGCTTACCACTGATCGAACCAAGGGGGCGGCTGCTGATG 540
 DB 481 GCCGCGCGGAGTTGGGCTTACCACTGATCGAACCAAGGGGGCGGCTGCTGATG 540

QY 541 GCGGCGGAGTGGCGGCGGCTGCGAACCGGCGGCTGCTGATGCGGCGGCGGCGGCG 600
 DB 541 GCGGCGGAGTGGCGGCGGCTGCGAACCGGCGGCTGCTGATGCGGCGGCGGCGGCG 600

QY 601 TACAACGAGCGCGGATGCGCAAGCGGAGCGGCGGAGCGGCTGATGATGATGATG 660
 DB 601 TACAACGAGCGCGGATGCGCAAGCGGAGCGGCGGAGCGGCTGATGATGATGATG 660

QY 661 ATGCAACAATCTGCGCAACTGACGCGGATTCGCGCGGAGTCCACACTGCTACTCA 720
 DB 661 ATGCAACAATCTGCGCAACTGACGCGGATTCGCGCGGAGTCCACACTGCTACTCA 720

QY 721 TCGGCGCTACGAGCTGAGAGGTGCGGCAAAAGTGGCGGAGTGTGATTTGGGGCGCTCTG 780
 DB 721 TCGGCGCTACGAGCTGAGAGGTGCGGCAAAAGTGGCGGAGTGTGATTTGGGGCGCTCTG 780

QY 781 GTGCGAGGCGGAGCGACCAATTAAGTCTGCAATTCAGTGTGCGGATATGAACCA 840
 DB 781 GTGCGAGGCGGAGCGACCAATTAAGTCTGCAATTCAGTGTGCGGATATGAACCA 840

QY 841 GGTGGGCTACTGTTGGATATGACCATCGACGAGGGGGGCTGTTTGAAGGCTACAGACCG 900
 DB 841 GGTGGGCTACTGTTGGATATGACCATCGACGAGGGGGGCTGTTTGAAGGCTACAGACCG 900

QY 901 ACCACTACGACACACCGACGCTTGGCGGTGACGACAGCAAGCTGTTTACTGCGTGGGAC 960
 DB 901 ACCACTACGACACACCGACGCTTGGCGGTGACGACAGCAAGCTGTTTACTGCGTGGGAC 960

QY 961 ATGGCGGCGTGGTGGCGAAGACGCTGACGCGTGCAGCAACGGGAGCATGCGGTAT 1020
 DB 961 ATGGCGGCGTGGTGGCGAAGACGCTGACGCGTGCAGCAACGGGAGCATGCGGTAT 1020

QY 1021 GTGCTGAGCTTGGCGACATGCGTGGGCGGCGGCGTGGCGGATTCGAGTCCGACTAGCC 1080
 DB 1021 GTGCTGAGCTTGGCGACATGCGTGGGCGGCGGCGTGGCGGATTCGAGTCCGACTAGCC 1080

QY 1081 AAAGCTCTTTCGACGACGAGGAGGCGGCTTACTGTCCGAACGGGTGGCCACGACTGGGG 1140
 DB 1081 AAAGCTCTTTCGACGACGAGGAGGCGGCTTACTGTCCGAACGGGTGGCCACGACTGGGG 1140

QY 1141 GTGCGCTTACGAGCGCGCGGCGGAGGCTGCTGCTGACTCTGCGCGCTCGTTAGCCGAG 1200
 DB 1141 GTGCGCTTACGAGCGCGCGGCGGAGGCTGCTGCTGACTCTGCGCGCTCGTTAGCCGAG 1200

QY 1201 C 1201
 DB 1201 C 1201

RESULT 12
 A87609 1235 bp DNA PAT 22-JUN-2000
 LOCUS Sequence 6 from Patent WO9836089.
 ACCESSION A87609
 VERSION A87609.1 GI:6736249
 KEYWORDS
 SOURCE unclassified.
 ORGANISM unclassified.
 REFERENCE 1 (bases 1 to 1235)
 Fliche, L. and Singh, M.
 TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
 JOURNAL Patent: WO 9836089-A 20-AUG-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES Location/Qualifiers
 source 1..1235
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 235 a 395 c 384 g 220 t 1 others
 ORIGIN

Query Match 98.8%; Score 1195; DB 5; Length 1235;
 Best Local Similarity 99.8%; Pred. No. 8.3e-164;
 Matches 1206; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATCTTGAGATTATCGAATCTTCTTCACTGAAAGCGTACAGTATCGAGAGGGTAATC 60
 DB 1 ATCTTGAGATTATCGAATCTTCTTCACTGAAAGCGTACAGTATCGAGAGGGTAATC 60

QY 61 ATGGCGGCGGATTCGACGAGACCAAAACGAATTCGGGGTGGCCATCGACCCCG 120
 DB 61 ATGGCGGCGGATTCGACGAGACCAAAACGAATTCGGGGTGGCCATCGACCCCG 120

QY 121 GCCGCGCTCGGGAACCTAACCCGTCGTGGCCATGAGTGTCTCATCGAGCAGTCCGGA 180
 DB 121 GCCGCGCTCGGGAACCTAACCCGTCGTGGCCATGAGTGTCTCATCGAGCAGTCCGGA 180

QY 181 GAGGCGCTCGGCTATCACGACGCGGATTCAGAGCGGCGGCAACTGGTCGGCACC 240
 DB 181 GAGGCGCTCGGCTATCACGACGCGGATTCAGAGCGGCGGCAACTGGTCGGCACC 240

QY 241 GCCGACGAGGTGGGCGGAGCGGATTTATGCTCAAGGTCAAGAACCGGATACGGCGG 300
 DB 241 GCCGACGAGGTGGGCGGAGCGGATTTATGCTCAAGGTCAAGAACCGGATACGGCGG 300

Oy	301	GAATACGCGCCGCTCGGACACCGGGGAGAACCTTGTCACGTTCTTGSCATTTGGCCGGCCMA	360
Db	301	GAATACGCGCCGCTCGGACACCGGGGAGAACCTTGTCACGTTCTTGSCATTTGGCCGGCCMA	359
Oy	361	CGTCTCTTGACCCGATGGGTGTGGATATCCGGCACCCACCTCAATTGGCCTAGAGACGGTC	420
Db	360	CGTCTCTTGACCCGATGGGTGTGGATATCCGGCACCCACCTCAATTGGCCTAGAGACGGTC	419
Oy	421	CAGACCGCCGACGGCGCATCTACCCCTTGCTTGCCCGGATGAGCGAAGTGGCCGGTCCGACTC	480
Db	420	CAGACCGCCGACGGCGCATCTACCCCTTGCTTGCCCGGATGAGCGAAGTGGCCGGTCCGACTC	479
Oy	481	GCCGCCGAGGTTGGCGGTTCTTCCACACTGATGCGAACCAGAGGGGGGGGGGGTGGTGTATG	540
Db	480	GCCGCCGAGGTTGGCGGTTCTTCCACACTGATGCGAACCAGAGGGGGGGGGGGTGGTGTATG	539
Oy	541	GCGGGGGTGGCCCGCTCGAACCGGCCGACGTCGTGTATGGCGGGCCGACCGCGCGGC	600
Db	540	GCGGGGGTGGCCCGCTCGAACCGGCCGACGTCGTGTATGGCGGGCCGACCGCGCGGC	599
Oy	601	TACAACGAGCGCCCATCGGCAACGGGATGGGGCGCACCGTTACGGTTCTAGACATCAAC	660
Db	600	TACAACGAGCGCCCATCGGCAACGGGATGGGGCGCACCGTTACGGTTCTAGACATCAAC	659
Oy	661	ATCGACAACTTGGCGCACTCGACGCCGAGTTCTGGGGCCGGATCCACACTCGCTACTCA	720
Db	660	ATCGACAACTTGGCGCACTCGACGCCGAGTTCTGGGGCCGGATCCACACTCGCTACTCA	719
Oy	721	TGCGCCACTGAGCTCGAGGGTGGCGTCAAACTGACCGACTGGTGAATTGGGGCCCTCTGTG	780
Db	720	TGCGCCACTGAGCTCGAGGGTGGCGTCAAACTGACCGACTGGTGAATTGGGGCCCTCTGTG	779
Oy	781	GTGCCAGCGCCCAANGCACCACAAATTAGTCTGCAATTCACTGTGGCGCATCTGAAACCA	840
Db	780	GTGCCAGCGCGCAANGCACCACAAATTAGTCTGCAATTCACTGTGGCGCATCTGAAACCA	839
Oy	841	GGTGGGTAAGTGGATATATAGCATGACACGAGCGGGCGTGTGCAAGGCTCACAGACGG	900
Db	840	GGTGGGTAAGTGGATATATAGCATGACACGAGCGGGCGTGTGCAAGGCTCACAGACGG	899
Oy	901	ACCACCTACGACCAACCCGACGTTGCGCGTGCACGACACGCGTGTATTACTGCGTGGCGAAC	960
Db	900	ACCACCTACGACCAACCCGACGTTGCGCGTGCACGACACGCGTGTATTACTGCGTGGCGAAC	959
Oy	961	ATGCCCGGCTGGGTGGCGAAGCTGCACTACGCGGTGACCAACCGGACGATGGCGCAT	1020
Db	960	ATGCCCGGCTGGGTGGCGAAGCTGCACTACGCGGTGACCAACCGGACGATGGCGCAT	1019
Oy	1021	GTGCTCGAGCTTGGCGACCATGAGTGGCGGGCGGCGTGGCGGTCGATCGGGCACTAGCC	1080
Db	1020	GTGCTCGAGCTTGGCGACCATGAGTGGCGGGCGGCGTGGCGGTCGATCGGGCACTAGCC	1079
Oy	1081	AAAGGCTTTTGGAGCGACAGAGGGGGGCTTACTGTCCGAAGGGGTGGCCACGACCTCGGGG	1140
Db	1080	AAAGGCTTTTGGAGCGACAGAGGGGGGCTTACTGTCCGAAGGGGTGGCCACGACCTCGGGG	1139
Oy	1141	GTGCCGTTACAGACGACCCCGGACGAGGTGCTGGCTGACTCTCGGCGCGCTGTTACGCCGAG	1200
Db	1140	GTGCCGTTACAGACGACCCCGGACGAGGTGCTGGCTGACTCTCGGCGCGCTGTTACGCCGAG	1199
Oy	1201	CNCACTGG 1209	
Db	1200	CACACGTGG 1208	

DEFINITION Sequence 8 from Patent WO9832862.
 ACCESSION A89751 GI:6738285
 VERSION A89751.1
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 1 (bases 1 to 1235)
 FLOHE, L. and Singh, N.
 L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
 Patent: WO 9832862-A 30-JUL-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 1. 1235
 Location/Qualifiers
 source
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 236 a 394 c 385 g 220 t
 ORIGIN

Query Match 98.8%; Score 1194.4; DB 5; Length 1235;
 Best Local Similarity 99.8%; Pred. No. 1e-163; 2; Indels 1; Gaps 1;
 Matches 1206; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATCTTGAGATTATCGAACTTTCTTACACTGAAGCGACAGTATCGAGAGGGGTAATC 60
 Db 1 ATCTTGAGATTATCGAACTTTCTTACACTGAAGCGTACAGTATCGAGAGGGGTAATC 60
 QY 61 ATGGGCTGGTATTCGACGAGACAAAACAGCAATTCGGGTGGGATCACCCG 120
 Db 61 ATGGGCTGGTATTCGACGAGACAAAACAGCAATTCGGGTGGGATCACCCG 120
 QY 121 GCCGCGCTCGGAGAACTAACCCGCTGCGCATGAGGTCTCATCCAGCAGGTGCCGA 180
 Db 121 GCCGCGCTCGGAGAACTAACCCGCTGCGCATGAGGTCTCATCCAGCAGGTGCCGA 180
 QY 181 GAGGCTCGGCTATCAGCGAGGAGTTTCAAGGCGGACGCGCAACTGTCGGGACC 240
 Db 181 GAGGCTCGGCTATCAGCGAGGAGTTTCAAGGCGGACGCGCAACTGTCGGGACC 240
 QY 241 GCCGACGAGTGTGGGCGGAGCGTGTATTTGCTCAAGGTCAAAAGACCGATAGCGGC 300
 Db 241 GCCGACGAGTGTGGGCGGAGCGTGTATTTGCTCAAGGTCAAAAGACCGATAGCGGC 300
 QY 301 GAATACGCGCGCTCGGACAGAGGCAATCTGTTCAGCTTTCGATTGGCGGCTCA 360
 Db 301 GAATACGCGCGCTCGGACAGAGGCAATCTGTTCAGCTTTCGATTGGCGGCTCA 360
 QY 361 CGTCTTGACACGATGGTGTGATTCGCGACACGTCATTTGCTACGAGACCGTC 420
 Db 361 CGTCTTGACACGATGGTGTGATTCGCGACACGTCATTTGCTACGAGACCGTC 420
 QY 421 CAGACCGCGGAGCGGACGCTACCTTGTGCGGATGAGCGAAGTCCCGGTGCACTC 480
 Db 421 CAGACCGCGGAGCGGACGCTACCTTGTGCGGATGAGCGAAGTCCCGGTGCACTC 480
 QY 481 GCCGCGGAGTGTGGGCTTACCACTGTGGAACCAAGGGGCGCGGTGTCTGATG 540
 Db 481 GCCGCGGAGTGTGGGCTTACCACTGTGGAACCAAGGGGCGCGGTGTCTGATG 540
 QY 541 GCGGCGGTGCGCGGCGTGAACCGCGGCGAGCTGTGTGATGCGCGCGGACCGCGGC 600
 Db 541 GCGGCGGTGCGCGGCGTGAACCGCGGCGAGCTGTGTGATGCGCGCGGACCGCGGC 600
 QY 601 TACAAAGCGCGGATGCGCAAGGCAATGGGCGGACCGTTAGGTTTGAACATCAAC 660
 Db 601 TACAAAGCGCGGATGCGCAAGGCAATGGGCGGACCGTTAGGTTTGAACATCAAC 660
 QY 661 ATGCAAACTTGGCAACTCGAGCGCGAGTTTGGCGCGGATCAACACTCGCTACTCA 720
 Db 661 ATGCAAACTTGGCAACTCGAGCGCGAGTTTGGCGCGGATCAACACTCGCTACTCA 720
 QY 721 TCGGCTTACGAGCTGAGAGGTGCGGTCAAACTGCGGACTGTGATTTGGGCGGCTCTG 780
 Db 721 TCGGCTTACGAGCTGAGAGGTGCGGTCAAACTGCGGACTGTGATTTGGGCGGCTCTG 780

Db 720 TCGGCTTACGAGCTGAGAGGTGCGGTCAAACTGCGGACTGTGATTTGGGCGGCTCTG 779
 QY 781 GTGCGAGCGCGCAAGGACCAATTAAGTCCGAATTTCACTTCTGCGGATATGAACCA 840
 Db 781 GTGCGAGCGCGCAAGGACCAATTAAGTCCGAATTTCACTTCTGCGGATATGAACCA 839
 QY 841 GGTGCGTACTGATGATATAGCATGACACGAGGCGGTGTTTCAAGAGCTCACAGCCG 900
 Db 841 GGTGCGTACTGATGATATAGCATGACACGAGGCGGTGTTTCAAGAGCTCACAGCCG 899
 QY 901 ACCACTACGACACCGCGAGCTTGGCGTACAGACAGCTGTTTACTGCTGGGGAAC 960
 Db 901 ACCACTACGACACCGCGAGCTTGGCGTACAGACAGCTGTTTACTGCTGGGGAAC 959
 QY 961 ATGCGCGCTGAGTGGCGAAGAGCTGACCTAGCGGCTGACCAAGCGAGATGCCGTAT 1020
 Db 961 ATGCGCGCTGAGTGGCGAAGAGCTGACCTAGCGGCTGACCAAGCGAGATGCCGTAT 1019
 QY 1021 GTGCTGAGCTTGGCGACCATGCTGCGGCGGCGGTGCGGATCCGCACTAGCC 1080
 Db 1021 GTGCTGAGCTTGGCGACCATGCTGCGGCGGCGGTGCGGATCCGCACTAGCC 1079
 QY 1081 AAAGCTTTGAGAGCAAGAGGCGGCTTACTGTCCGAAGGGGGGCGACGACTGGGG 1140
 Db 1081 AAAGCTTTGAGAGCAAGAGGCGGCTTACTGTCCGAAGGGGGGCGACGACTGGGG 1139
 QY 1141 GTGCGTTCACGAGCGCGCGGCGGCTGCGCTGACTGTGCGCGGCTGTACGCGGAG 1200
 Db 1141 GTGCGTTCACGAGCGCGCGGCGGCTGCGCTGACTGTGCGCGGCTGTACGCGGAG 1199
 QY 1201 CACACGTG 1209
 Db 1201 CACACGTG 1208

Search completed: June 22, 2000, 14:57:02
 Job time: 17837 sec

Result	Query	Match	Length	DB	ID	Description
No.	Score					
1	1192	98.6	1245	1	V49626	Mycobacterium tube
2	1192	98.6	1260	1	V49510	Mycobacterium sp.
3	1192	98.6	1260	1	V49625	Mycobacterium tube
4	467.6	38.7	682	1	V49511	Mycobacterium mari
5	277.2	22.9	1125	1	N91423	Sequence of heat-r
c	6	212.4	17.6	28171	V52145	Streptococcus pneu
5	212.4	17.6	9280	1	V74442	Staphylococcus aur
c	7	178.4	14.8	1074	T67971	H. pylori membrane
8	94.8	7.8	1074	1	V43039	Streptococcus pneu
c	9	69.2	5.7	544	V43039	Streptococcus aur
10	57.6	4.2	31096	1	V47370	Randomising oligon
c	11	50.8	4.2	390	Q21833	PCR primer for 5'
c	12	50.8	4.2	390	Q36859	Randomising oligon
13	50.4	4.2	1413	1	X34205	Mycobacterium spec
14	50.4	4.2	1722	1	X34206	Mycobacterium spec
15	50.4	4.2	15872	1	T88715	Streptomyces venez
16	49	4.1	1269	1	003065	Sequence encoding
c	17	47.6	3.9	3946	T93610	Mycobacterium tube
18	47.6	3.9	28598	1	T67679	Sorangium cellulos
19	47.6	3.9	28958	1	T89956	Forangium cellulos
20	47.6	3.9	49377	1	V05287	The soraphen biosy
c	21	46.4	3.8	985	V44439	Mycobacterium tube
c	22	46.4	3.8	985	V64548	M. tuberculosis
23	46	3.8	1620	1	Q2482	groEL-1 gene codin
24	46	3.8	2582	1	T3117	Actinoplanes sp. a
25	46	3.8	2668	1	Q22485	groEL-1 gene. Reco
26	46	3.8	17955	1	V66642	Actinoplanes sp. a
27	45.4	3.8	833	1	Q64203	snab gene encoding
28	45.4	3.8	5392	1	Q64201	Sequence compisin
29	42.4	3.5	1833	1	Q64206	snBR gene encoding
30	42.4	3.5	2185	1	V48066	Clone P5-4 encodin
31	42.4	3.5	2186	1	Q93465	S. lividans proteas
32	42.4	3.5	24739	1	T93095	Streptomyces freno
33	42.4	3.5	24779	1	V25925	Streptomyces roseo
34	42.2	3.5	329	1	V44425	Mycobacterium tube

C	35	42.2	3.5	329	1	V64534	M. tuberculosis Im
	36	42	3.5	3946	1	T93610	Mycobacterium tube
	37	42	3.5	4018	1	O63879	Polyhydroxyalkanoa
C	38	41.8	3.5	882	1	V44403	Mycobacterium tube
	39	41.8	3.5	882	1	V64512	M. tuberculosis Im
	40	41.6	3.4	12036	1	O04668	Fha structural gen
	41	41.2	3.4	2414	1	O05926	Sequence encoding
	42	41	3.4	29879	1	O46806	eryA region of S. l
	43	40.8	3.4	2151	1	O76252	HSV-2 protease, TC
C	44	40.8	3.4	2151	1	O76261	HSV-2 protease/ICP
C	45	40.8	3.4	2472	1	O84671	HSV-2 UL26 gene. N

ALIGNMENTS

RESULT	1
ID	V49626
AC	V49626 standard; DNA; 1245 BP.
DT	20-NOV-1998 (first entry)
DE	Mycobacterium tuberculosis L-Alanine dehydrogenase 40 KD antigen.
KW	ss; Alanine dehydrogenase; tuberculosis; TB; mycobacteria; L-alanine.
OS	Mycobacterium tuberculosis.
PN	MO9836089-A2.
PD	20-AUG-1998.
PF	29-JAN-1998; E00483.
PR	29-JAN-1997; EP-101338.
PA	(FLOH/) FLOHE L.
PI	Flohe L, Hutter B, Kolz A, Singh M;
DR	Wpi: 98-457123/39.
PT	Diagnosis of, e.g. tuberculosis from alanine dehydrogenase activity
PT	- useful for, e.g. for diagnosis, differentiation of strains,
PT	monitoring vaccination and identification of mycobacterial
PT	inhibitors
PS	Disclosure; Fig 3.19; 55pp; German.
CC	The Mycobacterium tuberculosis L-alanine dehydrogenase (AlaDH) is used in
CC	the production of kits for diagnosing tuberculosis (TB) and other
CC	mycobacterial infections in humans or animals. Kits are used for direct
CC	diagnosis of TB on clinical samples (e.g. body fluids) and can
CC	differentiate between pathogenic and non-virulent strains, e.g. for
CC	identifying pathogens of the M.tuberculosis (M. t.) complex. The kit may
CC	also be used to identify substances that inhibit mycobacteria, for
CC	combating epidemics and for vaccination follow-up. Oligonucleotides
CC	derived from AlaDH are used similarly in diagnostic hybridisation tests,
CC	also for culture confirmation of isolated strains and for chromosome
CC	fingerprinting to detect/differentiate between mycobacteria, and for
CC	L-alanine-specific biotransformation reactions. AlaDH is an early
CC	antigen, present extracellularly after only a few days of growth, making
CC	it an ideal drug target.
SQ	Sequence 1245 BP: 238 A: 398 C: 387 G: 222 T:

Query Match	Score	DB 1;	Length
98.68;	1192;		1245;

```

Best Local Similarity 99.48; Pred. NO. 2.0e-239;
Matches 1208; Conservative 0; Mismatches 1; Indels 6; Gaps 1.

```

Qy	1	ACCTTGCAGATTAATCAAGAACTTTCTCTCACTGGAAGCTACAGATGAGAGGGGGTATTC	60
Qy	1	ACCTTGCAGATTAATCAAGAACTTTCTCTCACTGGAAGCTACAGATGAGAGGGGGTATTC	60
Db	1	ACTCTTGCAGATTAATCAAGAACTTTCTCTCACTGGAAGCTACAGATGAGAGGGGGTATTC	60
Qy	61	ATGCGCGTGGTATTTCCGACGAGACCAAAAACAAG-----AATTCCGGGTGGCATC	114
Db	61	ATGCGCGTGGTATTTCCGACGAGACCAAAAACAAGATTCATTTCCGGGTGGCATC	120
Qy	115	ACCCGCGCGGGGTGGCGGAATTAACCGGTGGTGGCCATGAGGTGCATCATCAGGCAAGT	174
Db	121	ACCCGCGCGGGGTGGCGGAATTAACCGGTGGTGGCCATGAGGTGCATCATCAGGCAAGT	180
Qy	175	GCGGAGAGGGGTCCGGCTATCAACGACGGGATTTCAAGGGGGGCGGCGCCACTGGTC	234
Db	181	GCGGAGAGGGGTCCGGCTATCAACGACGGGATTTCAAGGGGGGCGGCGCCACTGGTC	240

QY 235 GGCACCGCCGACAGGTGTGGCCGACGCTGATTATTGCTCAAGTCGAAGAAGCGATA 294
 DB 241 GGCACCGCCGACAGGTGTGGCCGACGCTGATTATTGCTCAAGTCGAAGAAGCGATA 300
 QY 295 GCGGCGGAATACGCGCCCTGCGACAGCGGCGAGATCTTGTCAAGTCTTGCATTGGCC 354
 DB 301 GCGGCGGAATACGCGCCCTGCGACAGCGGCGAGATCTTGTCAAGTCTTGCATTGGCC 360
 QY 355 GCGTCACGTGCTTGCACCGATGCGTGTGGATTCCGCGACACGCTAATTGCTTACGAG 414
 DB 361 GCGTCACGTGCTTGCACCGATGCGTGTGGATTCCGCGACACGCTAATTGCTTACGAG 420
 QY 415 ACCGTTCAGACCGCCGACGCGGCGACCTACCCCTGCTTCCCGCATGACGAAAGTGGCCGT 474
 DB 421 ACCGTTCAGACCGCCGACGCGGCGACCTACCCCTGCTTCCCGCATGACGAAAGTGGCCGT 480
 QY 475 CGACTCGCCGCGCAGGTGTGGCGCTTACCACTGATGCGAACCAGAGGGGCGCGGTGTG 534
 DB 481 CGACTCGCCGCGCAGGTGTGGCGCTTACCACTGATGCGAACCAGAGGGGCGCGGTGTG 540
 QY 535 CTGATGGGCGGCGTGGCGCGCGGCGTGAACCGCGCGACGCGTGTGATCGCGCGCACCC 594
 DB 541 CTGATGGGCGGCGTGGCGCGCGGCGTGAACCGCGCGACGCGTGTGATCGCGCGCACCC 600
 QY 595 GCGGCGTACAAAGCAGCGCCGCGATGCGCAACGCGATGGCGCGGCGCTTACGCTTACAG 654
 DB 601 GCGGCGTACAAAGCAGCGCCGCGATGCGCAACGCGATGGCGCGGCGCTTACGCTTACAG 660
 QY 655 ATCAACATCGCAAACTTGGCGCACTGACGCGGCTGCGCGCGGATGACCA TCGC 714
 DB 661 ATCAACATCGCAAACTTGGCGCACTGACGCGGCTGCGCGCGGATGACCACTGCGC 720
 QY 715 TACATCATGCGCTACGAGCTGAGGCGTCCCTCAAACTGCGGACGCTGATGATGGGCGCC 774
 DB 721 TACATCATGCGCTACGAGCTGAGGCGTCCCTCAAACTGCGGACGCTGATGATGGGCGCC 780
 QY 775 GTCTGTGTCAGGCGCGCAAGGCGACCCAAATTAGTCTGGAATT ACTTGTGGCTATATG 834
 DB 781 GTCTGTGTCAGGCGCGCAAGGCGACCCAAATTAGTCTGGAATTACTTGTGGCGCATATG 840
 QY 835 AAACGAGTGGGCTACTGATGATATGATGATGATGATGATGATGATGATGATGATGATG 894
 DB 841 AAACGAGTGGGCTACTGATGATATGATGATGATGATGATGATGATGATGATGATGATG 900
 QY 895 CGACGACACCTTACGACACCCGACGCTTGGCGGTGACGACGACGCTTGTACTGCGTG 954
 DB 901 CGACGACACCTTACGACACCCGACGCTTGGCGGTGACGACGACGCTTGTACTGCGTG 960
 QY 955 GCGAATATGCGCGCTGCTGCTGCGAAGAGTGCAGCTACGCGGCTGACCAAGCGCGATG 1014
 DB 961 GCGAATATGCGCGCTGCTGCTGCGAAGAGTGCAGCTACGCGGCTGACCAAGCGCGATG 1020
 QY 1015 CCGTATGCTGACGCTTGGCGACCATGCTGGCGGCGCGGCTGCGGCTGCAATCGGCA 1074
 DB 1021 CCGTATGCTGACGCTTGGCGACCATGCTGGCGGCGCGGCTGCGGCTGCAATCGGCA 1080
 QY 1075 CTAGCCAAAGGCTTTCGACGACGACGAGGCGGCTTACTGCTGCAAGCGGCTGCGCACCGAC 1134
 DB 1081 CTAGCCAAAGGCTTTCGACGACGACGAGGCGGCTTACTGCTGCAAGCGGCTGCGCACCGAC 1140
 QY 1135 CTGGGGGTCGCTTACGAGCGCGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1194
 DB 1141 CTGGGGGTCGCTTACGAGCGCGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 QY 1195 GCGGAGCAGCTG 1209
 DB 1201 GCGGAGCAGCTG 1215

RESULT 2
 V49510
 ID V49510 standard; DN: 1260 BP.
 AC V49510;

DT 20-OCT-1998 (first entry)
 DE Mycobacterium sp. AladH DNA
 KW Alanine dehydrogenase; AladH; ADH; diagnosis; tuberculosis; pathogen;
 KW Swimmers disease; vaccine; epidemic; infection; identification; ss.
 OS Mycobacterium sp.
 PN WO9832862-A2.
 PD 30-JUL-1998.
 PF 29-JAN-1998: E00484.
 PR 29-JAN-1997: EP-101339.
 PA (FLOH/) FLOHE L.
 PI Flohe L, Hutter B, Kolk A, Singh M;
 DR WPI: 98-427958/36.
 PT Nucleic acid encoding alanine dehydrogenase of Mycobacterium marinum
 PT used for diagnosis of tuberculosis and other mycobacterial
 PT diseases, also for treatment and prevention, for drug screening and
 PT for bio-transformation
 PS Disclosure, Page 11; 57pp; German.
 CC This sequence encodes an alanine dehydrogenase (ADH) protein isolated
 CC from a Mycobacterium sp. This protein is used to diagnose tuberculosis
 CC and other mycobacterial infections (including 'swimmers' disease', caused
 CC by M. marinum, a fish pathogen) in humans or animals. The protein can
 CC also be used for control of epidemics and for vaccination, to screen for
 CC agents with anti-mycobacterial activity, and in bio-transformations that
 CC are specific for L-alanine. Also mycobacteria can be identified by
 CC analysis of genomic ADH sequences. ADH is an antigen that is secreted
 CC early during infection.
 SQ Sequence 1260 BP; 243 A; 403 C; 389 G; 225 T;

Query Match 98.6%; Score 1192; DB 1; Length 1260;

Best Local Similarity 99.4%; Pred. No. 2, 6e-239;

Matches 1208; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 1 ATCTGAGATTAATGAACTTCTTACACTGAAAGTATGAGAGGGGTATC 60
 DB 16 ATCTGAGATTAATGAACTTCTTACACTGAAAGTATGAGAGGGGTATC 75
 QY 61 ATGCGCGTGGTATTCGACCGAGACCAAAACACG-----AATTCGGGTGGCATC 114
 DB 76 ATGCGCGTGGTATTCGACCGAGACCAAAACACGAAATTCGAATTCGGGTGGCATC 135
 QY 115 ACCCGCGCGCGCGCGGGAACCTAACCCGCTGGGCGCATGAGTGTGCTATCCAGGAGGT 174
 DB 136 ACCCGCGCGCGCGCGGGAACCTAACCCGCTGGGCGCATGAGTGTGCTATCCAGGAGGT 195
 QY 175 GCGGAGAGGGCTGCGCTATACCGACCGCGGATTTCAAGGCGGCGGCGCAATGTGTC 234
 DB 196 GCGGAGAGGGCTGCGCTATACCGACCGCGGATTTCAAGGCGGCGGCGCAATGTGTC 255
 QY 235 GGCACCGCGCGACAGGTGGGCGCGAGCTGATTTAGTCTCAAGGTCGAAGAAGCGATA 294
 DB 256 GGCACCGCGCGACAGGTGGGCGCGAGCTGATTTAGTCTCAAGGTCGAAGAAGCGATA 315
 QY 295 GCGGCGGAATACGCGCCCTGCGACACGCGGCGAGATCTTGTTCAGCTTGTTCATTGGCC 354
 DB 316 GCGGCGGAATACGCGCCCTGCGACACGCGGCGAGATCTTGTTCAGCTTGTTCATTGGCC 375
 QY 355 GCGTCACGTGCTTGCACCATGCGTGTGATTCGCGACACGCTCAATTGCTTACGAG 414
 DB 376 GCGTCACGTGCTTGCACCATGCGTGTGATTCGCGACACGCTCAATTGCTTACGAG 435
 QY 415 ACCGTTCAGACCGCGCGAGCGCGACCTACCCCTGTTCCCGGATGAGCGAAGTGGCGGT 474
 DB 436 ACCGTTCAGACCGCGCGAGCGCGACCTACCCCTGTTCCCGGATGAGCGAAGTGGCGGT 495
 QY 475 CGACTCGCCGCGCAGGTGTGGCGCTTACCACTGATGCGAACCAGAGGGGCGCGGTGTG 534
 DB 496 CGACTCGCCGCGCAGGTGTGGCGCTTACCACTGATGCGAACCAGAGGGGCGCGGTGTG 555
 QY 535 CTGATGGGCGGCGTGGCGCGTGTGAACCGCGCGACGCTGCTGCTGATGCGCGCGGCGAC 594
 DB 556 CTGATGGGCGGCGTGGCGCGTGTGAACCGCGCGACGCTGCTGCTGATGCGCGCGGCGAC 615

Matches 604: Conservative 0: Mismatches 458: Indels 12: Gaps 4:

```

QY 61 ATGCGCGTGGTATTCGACCGACCAACAAACAAATTCGGGTGGCCATCACC 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 ATGAAGATCGGATTCACAAAGAAATCAAAAACAAATGAAGAACCGGCTCCCATC 60
QY 121 GCGGGGCGTGGGAAATACACCGTGTGGCATGAGTGGTCT--CATCGAGGAGTGC 177
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 GAGGCGGTATACGCTCTCAAGCGGGGCAATGAGTGTATGTGAGACGGAAGGC 120
QY 178 GAGAGGGGCTCGCTATACACCGAGGATTTCAAGGCGGAGCGCGCACTGTGTGC 237
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 GCGGCTGGGTGGGGTTCCTTCATTCGAGTATGAAGAACCGGGGCACTGTGTGC 180
QY 238 ACCGCGGACGAGTGTGGGCGACGCTGATTTATTTCTCAAGTCAAGAACCGATAGC 297
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 CGAAGCTGAGAGATGCTTGAAGCGGGGAGATGCTTTAAAGTGAAGAGCGCGTGC 240
QY 298 GCGGATATGCGCGCGCTGGACACGGGAGATCTTGTGACGTTCTGATTTGGCGGC 357
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 CGAGAGTTCGCTATTTTCCCGGATGATTTTGTACGATTTGCAATTTAGCCGCG 300
QY 358 TCACGTGCTTGCACCGAGTCCGTTGTGATTCGCGACCGACGTCATTCCTACGAGACC 417
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 GCGGAACGCGTCAAGAAAGCGTGCAGAGCAAAAAGTGGCGCATGCTTACGAGAGC 360
QY 418 GTCCAGACCGCGAGCGGCGCACTACCCCTGCTTCCCGCATGAGCAAGTCCCGGTGCA 477
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 GTACAGCTGCGCAAGCGGCTGCTGACACTGTGACCGCATGTGTAAGTCCCGGCGCC 420
QY 478 CTCGCGCGCCAGTGGTGGCTTACACACGATGCGCAACCAAGGGGGCGGGTGGTGC 537
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 421 ATGTGGTCAAGATGGCGCCCGAGTTTCTCGAAGAACCCCAAGGGGAAAGGCAATTTG 480
QY 538 ATGGCGGGGTGTCGCGGCTCGTCAACCGCGGAGCTGTGATCGCGCGCACCGCC 597
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 481 CTTCGGGGGTGCGCGGATGTTGGGGCGGCGCAAGTACATCATCGCGCGGCAAGGCG 540
QY 598 GCGTACACGACGCGCGCATGCGCAAGCGCATGCGCGCGCACCGTTCATACATC 657
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 541 GCGAGCAACGCGCGCAAAATCGGGTCTCGGGGCAACGCTGACGATTTTGGACAT 600
QY 658 AACATCGAATAAATTGCGAATCGAATCGCGCGGATTCGCGGATTCACACTGCTAC 717
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 601 AACGCGGACGCGTGCAGCTGATGTTGTGCGGACCACTGACGACGCTCATG 660
QY 718 TCATCGGCTTACGAGCTGAGGTCGCGCAACGTCGCGACCTGTGATGGGCGCTC 777
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 661 TTCACACTGTACATATATCCGAGTCCGAGTCCGCGAATTCGATTCGTCGTCGCTC 720
QY 778 CTGTGCGCAGGCGCCAGACCCCAATTAATGTCGAAATTCATCTGCGCATATGAAA 837
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 721 TTGATCCCGGGGGCGAA--AGCGAAGTGTGACGGAAGATGTTGCGCTCATGACG 777
QY 838 CGAGTGGGTACTGTGTGATATAGCCATCGACAGAGGGGGGCTGTTCCAA--GGCTCA 894
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 778 CCGGATTCGTTGTTGTTGTTAATGCGCATTTGACCAAGGCGGCAATTTTCAACGACGAC 837
QY 895 CGACGACCACTTACGACCAACCGAGTTCGCGCTGACAGACGCTGTTTACTGCTG 954
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 838 CGACGCTTACGACGACGATCGACATACGTCACAGCAGG.GTGTTCATCATCGCGCTC 17
QY 955 GCGAACAATGCGCGCTGAGTGGGAAAGATGAGTACGGCTGACGACCAAGAAAG 1014
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 898 GCGAACAATGCGCGC--GTGCGCGCAGCTGACATTCGCGCTTACGAACGTCACGATC 954
QY 1015 CGGATGCTGAGAGCTTGCACATGCTGCGGCGCGCGCTGCGCTGATCCGGA 1074
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 955 CGATACGCTTGAATATGCGCAAAAGGCTACCGCGGTTGCTTGAATACCGCGGG 1014
QY 1075 CTAGCCAAAGCTTTTCAGACGACGAGAGGGCGCTTACTGCGACGAGGTGCC 1128
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1015 CTGTTAAAGGATCAACACGCTCGACGCGGACATGCTAGACGAGCGGTGCGC 1068

```

```

RESULT 6
VS2155/c
ID V52155 standard; DNA: 28171 BP.
AC V52155;
DE 23-OCT-1998 (first entry)
KW Streptococcus pneumoniae genome fragment SEQ ID NO:22.
KW Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
OS computer readable medium; vaccine; pharmaceutical composition; ds.
PN Streptococcus pneumoniae.
PM 009818931-A2.
PD 07-MAY-1998.
PF 30-OCT-1997; U19588.
PR 31-OCT-1996; US-029960.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M,
PI Kunsch CA, Rosen CA.
DR WPI: 98-272225/24.
PT Computer readable medium with recorded Streptococcus pneumoniae
PT polynucleotide sequences - useful in diagnostic kits and assays, and
PT pharmaceutical compositions and vaccines for Streptococcus
PT pneumoniae
PS Claim 1: Page 273-289; 1409pp; English.
CC The present invention describes a computer readable medium which has
CC the nucleotide sequences SEQ ID NO:1 to 391 (V52134 to V52524) recorded,
CC on it, or a representative fragment or a sequence at least 95% identical
CC to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in SEQ ID NO:1
CC to 391 (V52134 to V52524) are genomic fragments from Streptococcus
CC pneumoniae. The present invention also describes an isolated nucleic acid
CC molecule encoding a homologue of any of the fragments of the S.pneumoniae
CC genome (SEQ ID NO:1 to 391) where the nucleic acid molecule is produced
CC by a process comprising: (a) screening a genomic DNA library using as a
CC probe a target sequence defined by any of the sequences in SEQ ID NO:1
CC to 391, identifying members of the library which contain sequences
CC that hybridise to the target sequence and isolating the nucleic acid
CC molecules from the members; or (b) isolating mRNA, DNA or cDNA produced
CC from an organism, amplifying nucleic acid molecules whose nucleotide
CC sequence is homologous to amplification primers derived from the
CC fragment of the S. pneumoniae genome to prime the amplification and
CC isolating the amplified sequences. The computer readable medium can be
CC used in a computer-based system for identifying fragments of the
CC S. pneumoniae genome of commercial importance, or expression modulating
CC fragments of the S. pneumoniae genome. Products from the present
CC invention can be used in diagnosis kits and assays, and pharmaceutical
CC compositions and vaccines for S. pneumoniae.
SQ Sequence 28171 BP; 8507 A; 5080 C; 5440 G; 8144 T;

Query Match 17.6%; Score 212.4; DB 1; Length 28171;
Best Local Similarity 53.9%; Pred. No. 6,5e-36;
Matches 533: Conservative 0: Mismatches 436: Indels 19: Gaps 4:
QY 61 ATGCGCGTGGTATTCGACCGACCAACAAACAAATTCGGGTGGCCATCACC 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21711 ATGTTAATCGGAATCCCAAAAATAAATAACGAACCGTGTGCGCTCACACT 21552
QY 121 GCGGCGTGCAGCACTACCGCTGTCGAGGAGTGTATCATCAGAGTGCAGG 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21651 GAGAGTGTTCATAGCTATGTTAGTGTGTCATCGTCTTATGGAAGAAATCTG 21592
QY 181 GAGGCGTGGCTATACCGACCGGATTTCAAGGCGGACCGCGCAACTGTCGCGCC 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21591 CTGGGTCTGAGTATGATGATGATCAAAAGCAAGGAGTGTGTGTGTGTGTGT 21932
QY 241 GCCGACCAAGTGTGGCGGACGCTGATTTATGCTCAAGGTCAAAGAACCATAGCGCG 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21531 GCTGTGGAAGCTTTGGG--ACACAGTGTGTTGGAAGTAAAGATCTTTAAGTCT 21475
QY 301 GAATACGGCGGCTTCGACGAGGCGCAATCTGTTACGCTTTCGATTTGGCGCGCTCA 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21474 GAATACGGTACTTGCAGACGATCTTCTCTTCCACTACTTGCATATGCGCGCTCT 21415

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 15:16:28 ; Search time 341.15 Seconds

(without alignments)
460.653 Million cell updates/sec

Title: US-09-362-485-9

Perfect score: 1209

Sequence: 1 ATCTTGCGAGTTATCGAC.....GTTACGCCGAGCNCACGTCG 1209

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

Issued Patents_NA:*
1: /cgn2_6/ptodata/2/1na/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/1na/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/1na/5C.COMB.seq:*
4: /cgn2_6/ptodata/2/1na/5D.COMB.seq:*
5: /cgn2_6/ptodata/2/1na/6.COMB.seq:*
6: /cgn2_6/ptodata/2/1na/PTCUS.COMB.seq:*
7: /cgn2_6/ptodata/2/1na/Backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47.6	3.9	1620	3 US-08-461-775-10	Sequence 10, Appl
2	47.6	3.9	28958	1 US-08-258-261B-6	Sequence 6, Appl
3	47.6	3.9	28958	1 US-08-456-837-6	Sequence 6, Appl
4	47.6	3.9	28958	1 US-08-457-342-6	Sequence 6, Appl
5	47.6	3.9	28958	1 US-08-457-342-6	Sequence 6, Appl
6	47.6	3.9	28958	2 US-08-458-076A-6	Sequence 6, Appl
7	47.6	3.9	28958	2 US-08-764-233A-4	Sequence 4, Appl
8	47.6	3.9	28958	2 US-08-457-335A-6	Sequence 6, Appl
9	47.6	3.9	28958	2 US-08-729-214-6	Sequence 1, Appl
10	47.6	3.9	49377	2 US-08-816-105A-2	Sequence 1, Appl
11	47.6	3.8	2582	4 US-08-816-105A-2	Sequence 2, Appl
12	45.4	3.8	833	3 US-08-403-852D-3	Sequence 3, Appl
13	45.4	3.8	5392	3 US-08-403-852D-3	Sequence 1, Appl
14	42.4	3.5	1833	3 US-08-403-852D-6	Sequence 3, Appl
15	42.4	3.5	2185	3 US-08-173-508-3	Sequence 3, Appl
16	42.4	3.5	2185	3 US-08-265-310-3	Sequence 3, Appl
17	41.2	3.4	2414	7 5248599-1	Patent No. 5248599
18	41.2	3.4	2668	3 US-08-461-775-11	Sequence 11, Appl
19	41.2	3.4	20235	2 US-07-642-734C-3	Sequence 3, Appl
20	40.8	3.4	44377	3 US-08-804-227C-7	Sequence 7, Appl
21	40.8	3.4	44377	4 US-08-804-198-1	Sequence 1, Appl
22	40.6	3.4	4257	3 US-08-690-473-1	Sequence 1, Appl
23	40.6	3.4	12001	2 US-08-458-568A-11	Sequence 11, Appl
24	40.2	3.3	2064	1 US-08-343-428-1	Sequence 1, Appl
25	40.2	3.3	11219	2 US-07-642-734C-1	Sequence 1, Appl
26	39.6	3.3	474	3 US-08-403-852D-14	Sequence 14, Appl
27	39.6	3.3	1524	7 5512669-1	Patent No. 5512669

28	39.6	3.3	1525	7 5229279-1	Patent No. 5229279
29	39.4	3.3	459	4 US-08-387-942C-35	Sequence 35, Appl
30	39.4	3.3	1998	7 5212296-8	Patent No. 5212296
31	39.4	3.3	12588	4 US-08-387-942C-1	Sequence 1, Appl
32	39.2	3.2	420	1 US-08-470-179-148	Sequence 148, App
33	39.2	3.2	2353	6 PCT-US92-06840-1	Sequence 1, Appl
34	39.2	3.2	43280	3 US-08-804-227C-1	Sequence 1, Appl
35	38.8	3.2	8051	4 US-08-576-626A-2	Sequence 2, Appl
36	38.8	3.2	11219	2 US-07-642-734C-1	Sequence 1, Appl
37	38.6	3.2	1187	2 US-08-440-856A-2	Sequence 2, Appl
38	38.6	3.2	3231	2 US-08-074-121-4	Sequence 4, Appl
39	38.6	3.2	3231	6 PCT-US94-06447-4	Sequence 4, Appl
40	38.4	3.2	1215	4 US-08-947-726A-1	Sequence 1, Appl
41	38.4	3.2	2048	5 US-08-776-251-1	Sequence 1, Appl
42	38.2	3.2	2109	4 US-08-555-568B-20	Sequence 20, Appl
43	38	3.1	30001	1 US-08-125-468-1	Sequence 1, Appl
44	38	3.1	30001	1 US-08-474-933-1	Sequence 1, Appl
45	37.8	3.1	2588	3 US-08-796-414B-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-461-775-10
; Sequence 10, Application US/08461775
; Patent No. 5858773
GENERAL INFORMATION:
APPLICANT: MAZODIER, Philippe
APPLICANT: GUGLIEMET, Gerard
TITLE OF INVENTION: REGULATORY NUCLEOTIDE SEQUENCE OF THE
TITLE OF INVENTION: INITIATION OF TRANSCRIPTION
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESS: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461.775
FILING DATE:
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/050.313
FILING DATE: 10-MAY-1993
APPLICATION NUMBER: FR 9011186
FILING DATE: 10-SEP-1990
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feury, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 010830-035
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1620 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1620
US-08-461-775-10

DB 25786 GCCTTCGTGCGAGGGGCTCTCTCCCTCGAGGAGCGGGGGCCCGCCATGCCCGCCCTGCGCAGC 25845
 QY 685 GCCGAGTTCTGCGGGCGGATCCACACTGCTACTGCGCTACGAGCTCGAGGGTGC 744
 DB 25846 AAGAGGCTACACCGCTGCGCGCAAGGGGGCCATGCGCGCTCGAGCTCGCGCTCC 25905
 QY 745 GTCAAAACGTGCGGACCTGCTGATTTGGGGGCGCTCTGCTGCTCGAGGGCC 792
 DB 25906 GACCTCCAGACCTACTGCTCTCTGCGGGGCGACAGGCTCTCATGCGC 25953

RESULT 3

US-08-456-837-6
 ; Sequence 6, Application US/08456837
 ; Patent No. 5643774

GENERAL INFORMATION:

APPLICANT: Schupp, Thomas
 APPLICANT: Ligon, James M.
 APPLICANT: Beck, James Joseph
 APPLICANT: Hill, Dwight Steven
 APPLICANT: Ryals, John Andrew
 APPLICANT: Gaffney, Thomas Deane
 APPLICANT: Lam, Stephen Ting
 APPLICANT: Hammer, Phillip E.
 APPLICANT: Uknes, Scott Joseph
 TITLE OF INVENTION: Genes for the synthesis of
 TITLE OF INVENTION: antipathogenic substances
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Ciba-Geigy Corporation
 STREET: 7 Skyline Drive
 CITY: Hawthorne
 STATE: NY
 COUNTRY: USA
 ZIP: 10532

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/456,837
 FILING DATE: 01-JUN-1995
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/457,205
 FILING DATE: 01-JUN-1995
 APPLICATION NUMBER: 08/258,261
 FILING DATE: 08-JUN-1994

ATTORNEY/AGENT INFORMATION:

NAME: Elmer, James Scott
 REGISTRATION NUMBER: 36,129
 REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8614
 TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
 LENGTH: 28958 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-456-837-6

Query Match 3.9%; Score 47.6; DB 1; Length 28958;
 Best Local Similarity 44.1%; Pred. No. 0.01;
 Matches 286; Conservative 0; Mismatches 359; Indels 3; Gaps 2;

QY 148 GGCATGAGTGTCTATCTCCAGGAGTGCAGAGGCGCTGCGCTATACCGACCGGAT 207
 DB 25306 GCCCAGAGACCTATACCTCTCGGATGTGGCTTATTTGGCTGGCCACACCGGGCCAC 25365
 QY 203 TTCAAGGGGGAGGCGCGCACTGCTGGCACCGCCGACGAGTGTGGCGGACGCTGAT 267
 DB 25366 TTGAGCACACGGGCGCTCTGTAGCC JACUUGCGAGAGAGTCTCTCGCGCTGAC 25425
 QY 268 TTATTGCTCAGGTCAAGAACCCATACGCGGGGAAATCGGCGGCTCTCGACAGGGCAG 327
 DB 25426 TCGCTGCGCCAGACMACCGCCCGGAGACCGCTCTCGAGAGGAGGAGGACAGGC 25485
 QY 328 ATCTGTTCAGCTTCTTGATTTGGCCGCTGACGCTGCTTGCACGATGCTGTGTGAT 387
 DB 25486 AAGCTGCTCTGCTTTCTTCTGGCAGAGCTGCGATGGGAAGGATGCGCTCTGCTG 25545
 QY 388 TCCGACACACGCTCAATTGCTTACGAGACCGTCCAGACCGCGGCGGACACTACCCCTG 447
 DB 25546 CTGACTCTCTGCGCGCTCTTCCGCGCTCAGCTCGAAGCATGGAGCGGCGCTGCTCT 25605
 QY 448 CTTCGCCGATGAGGAGAGTGGCGGTGACTTGGCCCGCAGGTTGGCGCTTACGAC--C 505
 DB 25606 CACGTGAGATGAGGCTGCTGCGCGCTCTGCGCGCGGAGGCGGCGCCCTCTCGAC 25665
 QY 506 TGATGCGAACCCAGGGGGCGCGGTGTGCTGATGGGCGGGTCCGCGGCTGAAACCG 565
 DB 25666 CCGCTGAGCTGTGACAGCCGCGCTCTTGGCTGATGCTCTCTGCGGCGCTCTG 25725
 QY 566 CCGAGC-TGCTGTGATGCGCGCGGCGGACCGCGGCTACAGCAGCGCGGATCGCAAC 624
 DB 25726 CCGTGGCTGGGCTGAGAGCCCGCGCGCTGCGGCGGACAGTCAAGGAGATCGCGCC 25785
 QY 625 GCGATGGGCGGACGCTTACGCTTACGATCAACATCGAACAACTTGGCAACTCGAC 684
 DB 25786 GCCTTCGTGCGAGGGCTCTCTCCCTCGAGGAGCGGGCGCGCGCTCTGCGCAGC 25845
 QY 685 GCCGAGTTCTGCGGGCGGATCCACACTGCTACTGCGCTACGAGCTCGAGGGTGC 744
 DB 25846 AAGAGGCTACACCGCTGCGCGCAAGGGGGCCATGCGCGCTCGAGCTCGCGCTCC 25905
 QY 745 GTCAAAACGTGCGGACCTGCTGATTTGGGGGCGCTCTGCTGCTCGAGGGCC 792
 DB 25906 GACCTCCAGACCTACTGCTCTCTGCGGGGCGACAGGCTCTCATGCGC 25953

RESULT 4

US-08-457-342-6
 ; Sequence 6, Application US/08457342
 ; Patent No. 5662898

GENERAL INFORMATION:

APPLICANT: Schupp, Thomas
 APPLICANT: Ligon, James M.
 APPLICANT: Beck, James Joseph
 APPLICANT: Hill, Dwight Steven
 APPLICANT: Ryals, John Andrew
 APPLICANT: Gaffney, Thomas Deane
 APPLICANT: Lam, Stephen Ting
 APPLICANT: Hammer, Phillip E.
 APPLICANT: Uknes, Scott Joseph
 TITLE OF INVENTION: Genes for the synthesis of
 TITLE OF INVENTION: antipathogenic substances
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Ciba-Geigy Corporation
 STREET: 7 Skyline Drive
 CITY: Hawthorne
 STATE: NY
 COUNTRY: USA
 ZIP: 10532

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,342
FILING DATE: 01-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-457-342-6

```

```

Query Match          3.9%; Score 47.6; DB 1; Length 28958;
Best Local Similarity 44.1%; Pred. No. 0.01;
Matches 286; Conservative 0; Mismatches 359; Indels 3; Gaps 2;

Oy 148 GCCCATGAGTCTCTATCCAGCAGGTGCGGAGAGGGCTCGCTATCCAGCAGCGGAT 207
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25306 GCCCAGCAGACCTACCTCCGCGATGCGCTATTGCTGCGCACCGCGCCGCAC 25345

Oy 208 TTCAAGCGCGGAGCGGCGCACTGTGCGGACCGCGGAGCGTGGGCGCAGCTGAT 267
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25366 TTGAGCAGACGGCGGCTCTGACCCACAGCGAGAGCTCTCGCGCTCGAC 25425

Oy 268 TTATGCTCAGGTAAAGAACCGATAGCGGCGGATAGCGCGCGCTCGACACGGGACG 327
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25426 TTGCTCGCGCAGCAAGCGCGCGCGGACCGCGCTCTCGGAGGAGGAGGACGGC 25485

Oy 328 ATCTGTTCAGCTTCTGCTATTTGGCCGGCTGACGTGCTTGCACGAGTGGTTGGAT 387
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25486 AAGCTGCTTCTGCTTTCTTCTGGCAAGGCTCGCAGTGGGAAG..ATGCCCTCTCGCTG 25545

Oy 388 TCCGCGACACGTCATATGCTTACGAGACCGTCCAGACCGCGGCGGACCTACCCCTG 447
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25546 CTCGACTCCTCCCGCTCTTCGCGCTCAGCTCGAAGCATGCGACGCGCGCTCGCTCT 25605

Oy 448 CTTGCGCGGATAGCGGAATGCGCGGCTGACATCGCGCGCGGAGTTGGCGCTTACGAC--C 505
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25606 CACGTCGAGTGGAGCGCTGCTCTGCTGCGCGCGGAGGAGGCGCGCGCTCTCGAC 25665

Oy 506 TGATCGGAACCAAGGGGCGCGGCTGATGAGCGGGGCTGCGCGCGTGAACCGG 565
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25666 CCGCTCGAGCTGCTACAGCGCGCGCTTTCGCGCTGATGCTCTCCCTGCGCGCGCTCTG 25725

Oy 566 CCGACG-TGCTGTGATGCGCGCGGACCGCGGCTTACACGACCGCGGATCGCCAC 624
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25726 CGCTGCTGCGGCTAGAGCGCGCGCGCTGCTGCGGACAGTCAAGGCGGAGTCCGCGC 25785

Oy 625 GGCATGGCGGCAACGCTTATAGCATCAATCAATCGAATAAACTTGGCAACTCGAC 684
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25786 GCTTCTGTCGAGCGGCTCTCTCCCTCGAGGAGCGCGCGCGGATCGCGCGCTTGGCAGC 25845

Oy 685 GCCGAGTCTGCGCGCGGATCCACACTGCTACTCATCGAGCTTACAGAGTTCGAGGCTGC 744
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25846 AAGGCGTACACAGCGTGGCGGCAAGCGGGGCGCATGCGCGCGCTGAGAGTTCGCGCGCTTC 25905

```

```

Oy 745 GTCAAGTCCGACCTGTGATTTGGGCGGCTCTCTGATGCCAGGCGCC 792
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25906 GACCTCCAGACCTACTCTCTCTCTGCGGCGGAGCGGCTCTCATCGCC 25953

RESULT 5
US-08-457-646A-6
Sequence 6, Application US/08457646A
Patent No. 5679560
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Liqon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John / drew
APPLICANT: Gaffney, Thomas Deane
APPLICANT: Lam, Stephen Ting
APPLICANT: Hammer, Phillip E.
APPLICANT: Uknes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,646A
FILING DATE: 01-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-457-646A-6

Query Match          3.9%; Score 47.6; DB 1; Length 28958;
Best Local Similarity 44.1%; Pred. No. 0.01;
Matches 286; Conservative 0; Mismatches 359; Indels 3; Gaps 2;

Oy 148 GCCCATGAGTCTCTATCCAGCAGGTGCGGAGAGGGCTCGCTATCCAGCAGCGGAT 207
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25306 GCCCAGCAGACCTACCTCCGCGATGCGCTATTGCTGCGCACCGCGCCGCAC 25365

Oy 208 TTCAAGCGCGGAGCGGCGCACTGTGCGGACCGCGGAGCGTGGGCGCAGCTGAT 267
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25366 TTGAGCAGACGGCGGCTCTGACCCACAGCGAGAGCTCTCGCGCTCGAC 25425

```

```

QY 268 TTATTGCTCAAGCTCAAGAACCGATAGCGGGGAATACGCCGCCCTCGACACGGCCAG 327
DB 25426 TCGCTCGCCCAAGAGAGCCGCCCGGAGACCGCTCTCCGAGAGGAGGAGGAGCCACGGC 25485
QY 328 ATCTGTTCACGTTCTTGATTTGGCCCGGTGACAGGCTTGACAGCGATGCGTTGGAT 387
DB 25486 AAGCTGCTTCTGCTTCTTCTGGGCAAGGCTGCGACATGGGAAAGGATGGCCCTTCGCTG 25545
QY 388 TCCGACACCACTCAATTCCTACGAGACCGTCCAGACCGCCGCGGCTACACCCCTG 447
DB 25546 CTCGACTCTCCGCCCTCTCCGCGCTGAGCTCGAAGCATGGAGAGGCGCGCTCGCTCT 25605
QY 448 CTTCGCCCGATAGCGAGTCCCGCGGTGACCTGCGCCCGGAGTTGGCGCTTACAC--C 505
DB 25606 CACGTGAGTGGAGCTGCTCCCGCTGCGCCCGGAGGCGCGCCCTCCCTCGAC 25665
QY 506 TCAATCGAACCAGAGGGGCGCGGTGCTGATGGGCGGGTCCCGGCGTGAACCGG 565
DB 25666 CCGCTCGAGCTGTACAGCGCCCGCTTTCGCCGCTATGGTCTCCCTGCGCGCTCTG 75725
QY 566 CCGAGC-TGCTGATGTCGCGCCCGGACCGCGGCTACAAJGACCGCCGATCGCCAAAC 624
DB 25726 CCTCGCTCGGGGTAGAGCCCGCGCGCTGCTGCGCACAGTCAGGCGGAGATCGCGCC 25785
QY 625 GGCATGGCGCGACCGTTACGGTTCTAGACATCAATGACAACTTGGCACTCGAC 684
DB 25786 GCGTTCGTGCGAGCGCTCTCTCCCTGAGAGAGCGCGCGCTCGCCCTCGCGCAC 25845
QY 685 GCGGAGTTCTGGCGCGGATCCACATTCCTACTCATCGGCGCTAGAGGTGCGGGGTGCG 744
DB 25846 AAGGCGCTACACCGCTCGCCGCAAGGGGCGCATGGCGCGCTCGAGCTCGCGCTCC 25905
QY 745 GTCAAACTGCGACCTGCTGATTGGGCGCTCTGTCGCAAGCGCC 792
DB 25906 GACCTCCAGACTACTCGCTCCCTGCGGGGAGACAGGCTCTCATCGCC 25953

```

RESULT 6

US-08-458-076A-6
Sequence 6, Application US/08458076A

GENERAL INFORMATION:

PATENT NO. 5698425
 APPLICANT: Schupp, Thomas
 APPLICANT: Ligon, James M.
 APPLICANT: Beck, James Joseph
 APPLICANT: Hill, Dwight Steven
 APPLICANT: Ryals, John Andrew
 APPLICANT: Gaffney, Thomas Deane
 APPLICANT: Lam, Stephen Ting
 APPLICANT: Hammer, Phillip E.
 APPLICANT: Unkes, Scott Joseph
 TITLE OF INVENTION: Genes for the synthesis of
 TITLE OF INVENTION: antipathogenic substances
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Ciba-Geigy Corporation
 STREET: 7 Skyline Drive
 CITY: Hawthorne
 STATE: NY
 COUNTRY: USA
 ZIP: 10532
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/458,076A
 FILING DATE: 01-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/457,205

FILING DATE: 01-JUN-1995
 APPLICATION NUMBER: 08/258,261
 FILING DATE: 08-JUN-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Elmer, James Scott
 REGISTRATION NUMBER: 36,129
 REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8614
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 28958 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 US-08-458-076A-6

Query Match

3.9%; Score 47.6; DB 2; Length 28958;

Best Local Similarity 44.1%; Pred. No. 0.01; Mismatches 359; Indels 3; Gaps 2;

```

QY 148 GGCATGAGTGTCTCATTCAGACAGTCCGAGAGGGCTCGCTATACCGACGGAT 207
DB 25306 GCCACGACGACCTCACCTCGCGATGTGCTTTCGTCGTCACACCGCGCCAC 25365
QY 208 TTCAAGGGCGGCGCGCCCAATGTGTCGCAACCGCGCAAGGTGGCGGACGGTAT 267
DB 25366 TTCGAGACACCGCGCTCTGTAAGCCACAAACCGCGAGACTCTCTCGCGTAC 25425
QY 268 TTATTGCTCAAGCTCAAGAACCGATAGCGGGGAATACGGCGCGCTCGACAGCGG 327
DB 25426 TCGCTCGCCCAAGAGAGCCGCCCGGAGACCGCTCTCCGAGAGGAGGAGGAGCCACGGC 25485
QY 328 ATCTGTTCACGTTCTTGATTTGGCCCGGTGACAGGCTTGACAGCGATGCGTTGGAT 387
DB 25486 AAGCTGCTTCTGCTTCTTCTGGGCAAGGCTGCGACATGGGAAAGGATGGCCCTTCGCTG 25545
QY 388 TCCGACACCACTCAATTCCTACGAGACCGTCCAGACCGCCGCGGCTACACCCCTG 447
DB 25546 CTCGACTCTCCGCCCTCTCCGCGCTGAGCTCGAAGCATGGGCGCGCTCGCTCT 25605
QY 448 CTTCGCCCGATAGCGAGTCCCGCGGTGACCTGCGCCCGGAGTTGGCGCTTACAC--C 505
DB 25606 CACGTGAGTGGAGCTGCTCCCGCTGAGAGAGCGCGCGCTCGCCCGGAGGCGCGCCCTCGCGAC 25665
QY 506 TCAATCGAACCAGAGGGGCGCGGTGCTGATGGGCGGGTCCCGGCGTGAACCGG 565
DB 25666 CCGCTCGAGCTGTACAGCGCCCGCTTTCGCCGCTATGGTCTCCCTGCGCGCTCTG 25725
QY 566 CCGAGC-TGCTGATGTCGCGCCCGGACCGCGGCTACAAJGACCGCCGATCGCCAAAC 624
DB 25726 CCTCGCTCGGGGTAGAGCCCGCGCGCTGCTGCGCACAGTCAGGCGGAGATCGCGCC 25785
QY 625 GGCATGGCGCGACCGTTACGGTTCTAGACATCAATGACAACTTGGCACTCGAC 684
DB 25786 GCGTTCGTGCGAGCGCTCTCTCCCTGAGAGAGCGCGCGCTCGCGCTCGCGCAC 25845
QY 685 GCGGAGTTCTGGCGCGGATCCACATTCCTACTCATCGGCGCTAGAGGTGCGGGGTGCG 744
DB 25846 AAGGCGCTACACCGCTCGCCGCAAGGGGCGCATGGCGCGCTCGAGCTCGCGCTCC 25905
QY 745 GTCAAACTGCGACCTGCTGATTGGGCGCTCTGTCGCAAGCGCC 792
DB 25906 GACCTCCAGACTACTCGCTCCCTGCGGGGAGACAGGCTCTCATCGCC 25953

```

RESULT 7

US-08-764-233A-4

```

: Sequence 4, Application US/08764233A
: Patent No. 5716849
:
: GENERAL INFORMATION:
: APPLICANT: Ligon, James M.
: APPLICANT: Schnupp, Thomas
: APPLICANT: Beck, James J.
: APPLICANT: Hill, Dwight S.
: APPLICANT: Neff, Suzanna
: APPLICANT: Ryals, John A.
: TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Ciba-Geigy Corporation
: STREET: 520 White Plains Road, P.O. Box 2005
: CITY: Tarrytown
: STATE: NY
: COUNTRY: USA
: ZIP: 10591
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/764,233A
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/729,214
: FILING DATE: 09-OCT-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/258,261
: FILING DATE: 08-JUN-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Melgs, J. Timothy
: REGISTRATION NUMBER: 38,241
: REFERENCE/DOCKET NUMBER: 1506/CIP6
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (919) 541-8587
: TELEFAX: (919) 541-8689
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 28958 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: ORIGINAL SOURCE:
: ORGANISM: Sorangium cellulosum
: IMMEDIATE SOURCE:
: CLONE: p98/1
:
: US-08-764-233A-4
:
: Query Match 3.9%; Score 47.6; DB 2; Length 28958;
: Best Local Similarity 44.1%; Pred. No. 0.01;
: Matches 286; Conservative 0; Mismatches 359; Indels 3; Gaps 2.
:
: QY 148 GGCATGAGTCCTATCCAGCAGAGTGC CGAGAGGCTTCGGCTATC-ACGACGGGAT 207
: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: Db 25306 GCCCAGCAGACCTCACCCTCCGAGATGCGCTATTCCCTGCCACCCACCGCCCCAC 25365
:
: QY 208 TTCAGGGGCGCGGCGGCAACTGTGTCGCGACCGCGGACAGAGTGAGGCGCAGCGTGAT 267
: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: Db 25366 TTCGAGCAACCGCGCGCTCTCTGTACCCACAAACCGGAGAGACTCTCTCCGCGCTCGAC 25425
: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: QY 268 TTATTGCTCAAGAGTCAAGAACCGATAGCGCGGGAATACGCGCGCTCGGACAGGGCAG 327
: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: Db 25426 TCGCTCGCCCGAAGAACGCCGCCCGGAGACACCTGCTCGG-ACGGAGCGAAGCCACGGC 25485
:
: QY 328 ATCTGTTCACGCTCTGTGATTTGGCGGCGGTACAGTGTGTG-ACGATGCGTGTGTGAT 387
: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: Db 25486 AAGCTGCTCTTCTGCTCTTTCTCTGGGCAAGGCTCGCAAGTGGGAAGGATGGCCCTCTGCTG 25545

```

QY	388	TCGGGACCCACGCAATTCCTACGACGACGCTCCAGACCCGGCCGACGACTACCCCTG	447
Db	25546	CTGCACTCCCTCGCCCTCTCTCCGCGCTACGCTCGACGATCGACGCGCGCTCGCTCT	25605
QY	448	CTTGCCCGCATGACGCACTCGCCGCGTGCAGCTCGCCGCGCAGGTTGGCGCTTACCAC--C	505
Db	25606	CACGTGAGAGGAGCGCTGCTCGCTCGCTCGCGCGCGACGAGGGGCGCCCTCGCTCGAC	25665
QY	506	TGATCGCAACCCAGAGGGGCGCGGTGTCTGATGGCGGGGTTCGCCGCTGCAACCGG	565
Db	25666	CGGTCGACGATGCTACAGCCCGCCCTCTTTCGCTGATGCTCTCCCTGGGGGCGCTCTTG	25725
QY	566	CGCAGC-TCGTGTGATCGGGCGCGGACCGCGGCTTACAAACGACGCGCATCGCCACAC	624
Db	25726	CGCTCGCTCGGGGTAGAGCGCGCGCGCTGTCTGCGCCACAGTCAAGGGCAGATCGCGCC	25785
QY	625	GGCATGGGCGCGCAGCTTACGTTCTAGACATCAACATCGACAACTTGGCAACTCGAC	684
Db	25786	GCCTTGTGTCGACAGGCGCTCTCTCCCTTCAGAGACGCGCGCCGATCGCGCGCTGCGTAC	25845
QY	685	GGCGAGTTCTGGCGGCGGATTCACACTCGCTACTCTATCGGCTTCAGAGCTCGAGGGTGC	744
Db	25846	AAAGCGCTACACACCGCTCGCGCGCAACGGGGCATGGCGCGCGCTCGAGCTCGGCGCTCC	25905
QY	745	GTCAAACGTGCGACCTGGTATTTGGGCGCTCTCGTGGCCAGGCGCG	792
Db	25906	GACCTCGACGACTACTCGCTCGCTCGGGGCGCACAGGCTTCTCATCGCC	25953
RESULT	8		
US-08-457-335A-6			
: Sequence 6, Application US/08457335A			
: Patent No. 5731759			
GENERAL INFORMATION:			
APPLICANT: Schnupp, Thomas			
APPLICANT: Ligon, James M.			
APPLICANT: Beck, James Joseph			
APPLICANT: Hill, Dwight Steven			
APPLICANT: Rivals, John Andrew			
APPLICANT: Galfney, Thomas Deane			
APPLICANT: Lam, Stephen Ting			
APPLICANT: Hammer, Phillip E.			
APPLICANT: Uknes, Scott Joseph			
TITLE OF INVENTION: Genes for the synthesis of			
TITLE OF INVENTION: antipathogenic substances			
NUMBER OF SEQUENCES: 22			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Ciba-Geigy Corporation			
STREET: 7 Skyline Drive			
CITY: Hawthorne			
STATE: NY			
COUNTRY: USA			
ZIP: 10512			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: Patent Release #1.0, Version #1.25			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/457,335A			
FILING DATE: 01-JUN-1995			
CLASSIFICATION: 800			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: US 08/457,205			
FILING DATE: 01-JUN-1995			
APPLICATION NUMBER: 08/258,261			
FILING DATE: 08-Jun-1994			
ATTORNEY/AGENT INFORMATION:			
NAME: Elmer, James Scott			
REGISTRATION NUMBER: 36,129			
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3			
TELECOMMUNICATION INFORMATION:			

TELEPHONE: 919-541-8614
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 28958 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 US-08-457-335A-6

Query Match 3.9%; Score 47.6; DB 2; Length 28958;
 Best Local Similarity 44.1%; Pred. No. 0.01;
 Matches 286; Conservative 0; Mismatches 359; Indels 3; Gaps 2;

148 GGCATGAGTGTCTATCCAGGAGTGGCGGAGAGGGCTCGCTATCAACGAGGGGAT 207
 Db 25306 GCCACGACGACCTACCTCGGGATGGCTTATTCCTG6CCACCCGCGCCAC 25365
 QY 208 TTCAAGGCGC.AGGCGCGCACTGTGCGACCGCCGAC.AGATGTGGCCGACGCTGAT 267
 Db 25366 TTGAGGACCGGCGCGCTCTGTAAGCCCAACCGGAGAGTCTCTCCGGGCTCGAC 25425
 QY 268 TTATGCTCAGGTCAAGAACCGATAGCGCGGAATACGCGCGCTCGACACGGGCGAG 327
 Db 25426 TCGCTCGCCAGGACACGCGCCCGGACACCGCTCTCGAGAGGAGGAAAGCCAGCGC 25485
 QY 328 ATCTGTTCAGCTTCTTGTGATTTGGCGGCTCAGCTGCTTGCACGATGCTTGTGAT 387
 Db 25486 AAGCTGTGCTTGTCTTCTTGGGCAAGGCTCGACAGTGGGA.AGGATGACCTCTGCTG 25545
 QY 388 TCCGACACGCTCAATGCTACGAGACGCTCCAGACCGCGGACGCGCCTACCTG 447
 Db 25546 CTCGACTCCTCGCCCTCTTCGGCGCTCAGCTGGAAGATGGAGGCGCGCTCGCT 25605
 QY 448 CTTGCCCCATGACGAGTCCGCGCTGACACTGCGCCCGCAGGTTGGCTTACAC--C 505
 Db 25606 CAGCTGAGTGGAGCTGCTGCGCGCTGCGCGCGGAGAGGCGCCCTCTCGAC 25665
 QY 506 TGTGGAACCCAGAGGGGCGGCTGTGTATGAGGCGGGGCGCGGCTGGAACCGG 565
 Db 25666 CCGCTGAGCTGTACAGACCGCGCTTTCCTTTCCTATGCTCTCTCGGCGCTCTG 25725
 QY 566 CCGAGC-TGCTGTATTCGCGCGCGGACCGCGCTACAGCAGCAGCGCATGCGCAAC 624
 Db 25726 CGCTGCTGGGGTAAAGCGCGCGCGCTGTGCGGACAGTAGGGGCGAGATCGCGCG 25785
 QY 625 GGCATGGCGGACGCTTACGCTTACAGATCAGATGACAACTTGGGCACTGAC 684
 Db 25786 GCGTTCGTGAGGAGGCTCTCTCCCTGAGAGAGCGCGCGCTCGCGCTCGACG 25845
 QY 685 GCGAGTTCGCGGCGGATCCAGCTGCTACTATCGGCTACAGGCTGAGGGTGGC 744
 Db 25846 AAGGCTACACACCTCGCGCGGCAAGGGGCGCATGGCGCGCTCGACGCTCGCTCC 25905
 QY 745 GTCAAAAGTCCGACCTGTGATTTGGGCGCGCTGCTGTGCTCGACGCGC 792
 Db 25906 GACCTTCAGACCTACCTCTCTCGGCGGAGAGGCTCTCATCGC 25953

RESULT 9

US-08-729-214-6
 Sequence 6, Application US/08729214
 Patent No. 5817502
 GENERAL INFORMATION:
 APPLICANT: Ligon, James M.
 APPLICANT: Hill, Dwight Steven
 APPLICANT: Ryals, John Andrew
 APPLICANT: Hammer, Phillip E.
 APPLICANT: van Pee, Karl-Heinz

APPLICANT: Kliner, Sabine
 TITLE OF INVENTION: Genes for the synthesis of
 TITLE OF INVENTION: antipathogenic substances
 NUMBER OF SEQUENCES: 27
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Ciba-Geigy Corporation
 STREET: 520 White Plains Road
 CITY: Tarrytown
 STATE: NY
 COUNTRY: USA
 ZIP: 10591
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/729,214
 FILING DATE: TBA

CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Meigs, J. Timothy
 REGISTRATION NUMBER: 38,241
 REFERENCE/DOCKET NUMBER: CGC 1506/CIP5
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8587
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 28958 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 US-08-729-214-6

Query Match 3.9%; Score 47.6; DB 2; Length 28958;
 Best Local Similarity 44.1%; Pred. No. 0.01;
 Matches 286; Conservative 0; Mismatches 359; Indels 3; Gaps 2;

148 GGCATGAGTGTCTATCCAGGAGTGGCGGAGAGGGCTCGCTATCAACGAGGGGAT 207
 Db 25306 GCCACGACGACCTACCTCGGGATGGCTTATTCCTG6CCACCCGCGCCAC 25365
 QY 208 TTCAAGGCGCGAGCGCGCACTGTGCGACCGCCGAC.AGATGTGGCCGACGCTGAT 267
 Db 25366 TTGAGGACCGGCGCGCTCTGTAAGCCCAACCGGAGAGTCTCTCCGGGCTCGAC 25425
 QY 268 TTATGCTCAGGTCAAGAACCGATAGCGCGGAATACGCGCGCTCGACACGGGCGAG 327
 Db 25426 TCGCTCGCCAGGACACGCGCCCGGACACCGCTCTCGAGAGGAGGAAAGCCAGCGC 25485
 QY 328 ATCTGTTCAGCTTCTTGTGATTTGGCGGCTCAGCTGCTTGCACGATGCTTGTGAT 387
 Db 25486 AAGCTGTGCTTGTCTTCTTGGGCAAGGCTGCGAGTGGGAAGGATGACCTCTCGCTG 25545
 QY 388 TCCGACACGCTCAATGCTACGAGACGCTCCAGACCGCGGACGCGCCTACCTG 447
 Db 25546 CTCGACTCCTCGCCGCTCTTCGGCGCTCAGCTGGAAGATGGAGCGGCGCTCGCTCT 25605
 QY 448 CTTGCCCCATGACGAGTCCGCGCTGACACTGCGCCCGCAGGTTGGGCTTACAC--C 505
 Db 25606 CAGCTGAGTGGAGCTGCTGCGCGCTGCGCGCGGAGAGGCGCCCTCTCGAC 25665
 QY 506 TGTGGAACCCAGAGGGGCGGCTGTGTATGAGGCGGGGCTGCGCGCGTGAACGG 565
 Db 25666 CCGCTGAGCTGTACAGCGCGCGCTTTCCTGCGTATGCTCTCTCGGCGCTCTG 25725
 QY 566 CCGAGC-TGCTGTATGAGGCGCGCGGACCGCGGCTACAGCGAGCGCATGCGCAAC 624

```

Db 25726 CGCTGCTGCGGCTAGAGCGCGCGCGGCGGCGCAAGTCAGAGGCGAGATCGCGCG 25785
Oy 625 GGCATGCGCGCGACCGCTTACGTTCTAGACATCAACATCGACAACTCGGCACTCGAC 684
Db 25786 GCGTTGCTGCGAGGCGCTCTCTCCCTCGAGGAGCGGCGCGCGCGCGCTGCGCAGC 25845
Oy 685 GCGGAGTTGCGGCGCGATCCACACTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 744
Db 25846 AAGCGCTTACCAACCGTCCCGCGCAAGCGGCGCGCGCGCGCTGCGCGCGCTGCTG 25905
Oy 745 GTCAACGTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 792
Db 25906 GACCTTCAGACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25953

RESULT 10
US-08-764-233A-1
; Sequence 1, Application US/08764233A
; Patent No. 5716849
; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Schupp, Thomas
; APPLICANT: Beck, James J.
; APPLICANT: Hill, Dwight S.
; APPLICANT: Neff, Snezana
; APPLICANT: Ryals, John A.
; TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 520 White Plains Road, P.O. Box 2005
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/764,233A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/729,214
; FILING DATE: 09-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/258,261
; FILING DATE: 08-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: 1506/CIP6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49377 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Sorangium cellulosum
; IMMEDIATE SOURCE:
; CLONE: p98/1, pJL3, and pVKM15
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 383..760
; OTHER INFORMATION: /note- "Sorr"
; OTHER INFORMATION: /note- "This gene encodes a protein that is highly homologous

```

```

; OTHER INFORMATION: the reductase domains of type I PKs such as eryA from
; OTHER INFORMATION: Saccharopolyspora erythraea."
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 927..19874
; OTHER INFORMATION: /product= "Sora"
; OTHER INFORMATION: /note- "Gene product is highly homologous to type I PKs."
; OTHER INFORMATION: are known to be involved in the synthesis of polyketide
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 942..7115
; OTHER INFORMATION: /product= "Module 1 of Sora"
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 7203..12884
; OTHER INFORMATION: /product= "Module 2 of Sora"
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 13455..19616
; OTHER INFORMATION: /product= "Module 3 of Sora"
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 19871..45318
; OTHER INFORMATION: /product= "Sorb"
; OTHER INFORMATION: /note- "Gene product is highly homologous to type I PKs ge
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 19870..24556
; OTHER INFORMATION: /product= "Module 1 of Sorb"
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 24638..30820
; OTHER INFORMATION: /product= "Module 2 of Sorb"
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 30881..3544
; OTHER INFORMATION: /product= "Module 3 of Sorb"
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 35528..40114
; OTHER INFORMATION: /product= "Module 4 of Sorb"
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 40190..46318
; OTHER INFORMATION: /product= "Module 5 of Sorb"
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 46851..47891
; OTHER INFORMATION: /product= "Sorr"
; OTHER INFORMATION: /note- "The protein encoded by the sorr gene is highly
; OTHER INFORMATION: homologous to the methyltransferase from Streptomyces
; OTHER INFORMATION: hygrosopicus that is involved in the synthesis of the
; OTHER INFORMATION: polyketide rapamycin."
US-08-764-233A-1

Query Match 3.9%; Score 47.6; DB 2; Length 49377;
Best Local Similarity 44.1%; Pred. No. 0.012; Mismatches 339; Indels 3; Gaps 2;
Matches 286; Conservative 0; Mismatches 339; Indels 3; Gaps 2;

Oy 148 GGCATGAGTCTGATCCAGCGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 207
Db 41609 GCCACAGACGACCTCACCCTCGGATGCGCTATGCTGCTGCTGCTGCTGCTGCTGCTG 41668
Oy 208 TTCAGAGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTAT 267
Db 41669 TTCAGACACCGCGCGCTCTCTGACCCCAACCGCGGACGCTCTCTCTGCGCTGAC 41728
Oy 268 TTATGCTCAAGTCAAAAGACGATAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 327
Db 41729 TCGCTGCGCGCGAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 41788

```

QY 328 AACTGTTACGCTTCTTGATTTGGCCGGTCACGCTTGACACCGATGCTTGAT 387
 DB 41789 AACCTGCT 41848
 QY 388 TCCGACACCACTGCTTACGACACCGCTTACACCGCCGACGCTTACCTG 447
 DB 41849 CTCGACCT 41908
 QY 448 CTGCT 505
 DB 41909 CACGCTGAGTGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 41968
 QY 506 TGATGCAACCAAGGAGGCGCGGCTGCTGATGGCGGGTCCCGGCTGAA 565
 DB 41969 CGGCTGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 42028
 QY 566 CCGACG-TCGCTGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 624
 DB 42029 CGCT 42088
 QY 625 GGGATGGGCGGACCGCTTACGCTTCTGACATCAATCAAACTTCCGAACTG 684
 DB 42089 GCT 42148
 QY 685 GCGGAGTCTGCGCGCGGATTCACACTCTCTCTCTCTCTCTCTCTCT 744
 DB 42149 AAGCGCT 42208
 QY 745 GTCAAGTCT 792
 DB 42209 GACCT 42256

RESULT 11

US-08-816-105A-2
 ; Sequence 2, Application US/08816105A
 ; Patent No. 598982
 ; GENERAL INFORMATION:
 ; APPLICANT: Criegee, Anneliese, Dellweg, Hans-Georg; Lenz,
 ; APPLICANT: Jurgen, Schoder, Werner; Pape, Hermann;
 ; APPLICANT: Goetz, Klaus; Schaper, Beate; Hemker, Michael;
 ; APPLICANT: Piepersberg, Wolfgang; Distler, Jurgen;
 ; APPLICANT: Straumann, Ansgar
 ; TITLE OF INVENTION: PROCESSES FOR PREPARING ACARVITOL
 ; TITLE OF INVENTION: TRANSFERASE AND FOR USING IT IN THE
 ; TITLE OF INVENTION: CONVERSION OF ACARBOSE HOMOLOGUES
 ; TITLE OF INVENTION: INFO ACARBOSE, FOR THE PREPARATION
 ; TITLE OF INVENTION: OF ACARBOSE HOMOLOGUES
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Sprung Kramer Schaefer & Briscoe
 ; STREET: 660 White Plains Road
 ; CITY: Tarrytown
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10591-5144
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
 ; COMPUTER: Apple Macintosh 6500
 ; OPERATING SYSTEM: System 7.5
 ; SOFTWARE: WordPerfect 3.5
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/816,105A
 ; FILING DATE: 14-MAR-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: DE 19625269.5
 ; FILING DATE: 25-JUN-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: DE 19611252.4
 ; FILING DATE: 22-MAR-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kurt G. Briscoe

REGISTRATION NUMBER: 33,141
 REFERENCE/DOCKET NUMBER: Bayer 9814-KGB
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (914) 332-1700
 TELEFAX: (914) 332-1844
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2582 nucleotides
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-816-105A-2

Query Match 3.8% Score 46; DB 4: Length 2582;
 Best Local Similarity 47.3%; Pred. No. 0.013;
 Matches 139; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

QY 353 CCGCTGACGCTCTCTGACCGATGCTTGTGATTCGGGACACGCTCATGCTTACG 412
 DB 1576 CCGCTGCT 1635
 QY 413 AGACGCTCAGACCGCGGACGCGGACGCTTACGCTCTCTCTCTCTCTCTCT 472
 DB 1636 TGACGCTCAGACCGCGGACGCGGACGCTTACGCTCTCTCTCTCTCTCTCT 1695
 QY 473 GTGACT 532
 DB 1696 GCGAGTCT 1755
 QY 533 TGTGATGGGCGGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 592
 DB 1756 TCGGCGGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1815
 QY 593 CGCGCGCTTACAAAGCGCGGATCGCCATCGCCATCGGCGGCGGACCGTTACG 646
 DB 1816 CCGCGGAAACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1869

RESULT 12

US-08-403-852D-3
 ; Sequence 3, Application US/08403852D
 ; Patent No. 5891695
 ; GENERAL INFORMATION:
 ; APPLICANT: Blanc, Veronique
 ; APPLICANT: Blanchet, Francis
 ; APPLICANT: Crouzet, Joel
 ; APPLICANT: Jacques, Nathalie
 ; APPLICANT: Lacroix, Patricia
 ; APPLICANT: Thibaut, Denis
 ; APPLICANT: Zagorec, Monique
 ; APPLICANT: Debussche, Laurent
 ; TITLE OF INVENTION: Polypeptides Involved in the
 ; TITLE OF INVENTION: Biosynthesis of Streptogramins, Nucleotide Sequences
 ; TITLE OF INVENTION: Coding For These Polypeptides And Their Use
 ; NUMBER OF SEQUENCES: 43
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
 ; STREET: 1300 I Street, N.W., Suite 700
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005-3315
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/403,852D
 ; FILING DATE: 10-MAY-1995
 ; PRIOR APPLICATION DATA:

APPLICANT: Crouzet, Joel
APPLICANT: Jacques, Nathalie
APPLICANT: Lacroix, Patricia
APPLICANT: Thibaut, Denis
APPLICANT: Zagorec, Monique
APPLICANT: Debussche, Laurent
APPLICANT: De Crecy-Lagard, Valerie
TITLE OF INVENTION: Polypeptides Involved In The
Biosynthesis Of Streptogramins, Nucleotide Sequences
TITLE OF INVENTION: Biosynthesis Of Streptogramins And Their Use
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farbow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/403,852D
FILING DATE: 10-MAY-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/FR 93/00923
FILING DATE: 25-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/11441
FILING DATE: 25-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03806, 0054-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1833 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: S.pristinaespiralis
FEATURE:
NAME/KEY: CDS
LOCATION: 103..1689
US-08-403-852D-6

Query Match 3.5%; Score 42.4; DB 3; Length 1833;
Best Local Similarity 50.5%; Pred. No. 0.087;
Matches 103; Conservative 0; Mismatches 01; Indels 0; Gaps 0;
DB 533 TCTGTATGGCGGGGTGCGCGGCGAAGCGGCGTGTGTGATGCGGCGCGCA 592
413 TGTATGCTGCGAGTGGGTCTTGTGCGCGCGGCGTGTGCGCGCGTGTGCGCGAGACT 472
QY 593 CGCGCGGCTACAGAGCGCGCGCGATGCCAAGCGATGGCGCGGCGCGTGTAG 652
473 CGCGCGAAGTATCGCGCGCGCGCGCGATGGCGCGGCGCGTGTGCGCGACCA 532
DB 653 ACATGACATGCGAAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 712
533 CCACCGTCCCGCGATGACGAGTGTGCGAGCGACGACGACGACGACGACGACGACG 592
QY 713 GCTACTGATCGCGCGCTACGAGCTCG 736

DB 593 TGTGGGAGCGCGCGCGCTGACTGCG 616
RESULT 15
US-08-173-508-3
Sequence 3, Application US/08173508
Patent No. 5616485
GENERAL INFORMATION:
APPLICANT: Bartfield, Daniel
APPLICANT: Butler, Michael J.
APPLICANT: Hadary, Dany
APPLICANT: Jenish, David
APPLICANT: Krieger, Timothy
TITLE OF INVENTION: STREPTOMYCIN PROTEASES AND IMPROVED
STREPTOMYCIN STRAINS FOR EXPRESSION OF PEPTIDES AND
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/173,508
FILING DATE: 23-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 18740/125/CACO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2185 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 531..2069
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 531..902
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 903..2069
FEATURE:
NAME/KEY: misc_feature
LOCATION: 531..533
OTHER INFORMATION: /note="Met at position -124
represents fmet"
US-08-173-508-3

Query Match 3.5%; Score 42.4; DB 1; Length 2185;
Best Local Similarity 47.4%; Pred. No. 0.091;
Matches 127; Conservative 0; Mismatches 141; Indels 0; Gaps 0;
DB 525 CGCGGTGCTGATGGCGGGGTGCGCGCGTGTGCGCGCGCGCGCGCGCGCGCG 584
1250 CGCGGTGCTGATGGCGGGGTGCGCGCGTGTGCGCGCGCGCGCGCGCGCGCG 1309
QY 585 CGCGGACCGCGCGCTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 644

Db 1310 CCCCACGGCTTCTCTACACCGAGCCCGTGTCTGCGGCTTCATGTGGCGCCGAGCA 1369
 QY 645 GGTCTAGACATCAACATCGACAAACTTCGGCACTGACGCCGAGTTCTGCGCGCGAT 704
 Db 1370 CGGCGTCGACGTGACCAACAGCTATTACACCGACCCGTTGTTACTTCAACTGCAAGGA 1429
 QY 705 CCACACTGGCTACTCATGCGGCTAGAGCTGAGGGTGCCTCAAAAGTGGCGACTGT 764
 Db 1430 CGACCCGACCGAAGAGCGCTGTCGAGGCCCTCTCGCGGCGCTCCGTTAGCGGAGAA 1489
 QY 765 GATTGGGCGCTCTGTCGTCAGCGGCC 792
 Db 1490 GAAGGGCGGGTCAACGTGCGCGGCC 1517

Search completed: June 22, 2000, 15:18:59
 Job time: 18617 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 12:07:56 ; Search time: 5541.94 Seconds
(without alignments)
884.230 Million cell updates/sec

Title: US-09-362-485-9
Perfect score: 1209
Sequence: 1 ATCTGCGAGTATATCGAAC.....GTTACGCCGAGCNCACGTGG 1209

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 202611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: em_est20:*
21: em_est21:*
22: em_est22:*
23: em_est23:*
24: em_est24:*
25: em_est25:*
26: em_est26:*
27: em_est27:*
28: em_est28:*
29: em_est29:*
30: em_est30:*
31: em_est31:*
32: em_est32:*
33: em_est33:*
34: em_est34:*
35: em_est35:*
36: em_est36:*
37: em_est37:*
38: em_est38:*
39: em_est39:*
40: em_est40:*
41: em_est41:*
42: em_est42:*
43: em_est43:*
44: em_est44:*
45: em_est45:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: gb_est33:*
53: gb_est34:*
54: gb_est35:*
55: gb_est36:*
56: gb_est37:*
57: gb_est38:*
58: gb_est39:*
59: gb_est40:*
60: gb_est41:*
61: gb_est42:*
62: gb_est43:*
63: gb_est44:*
64: gb_est45:*
65: gb_est46:*
66: gb_est47:*
67: gb_est48:*
68: gb_est49:*
69: gb_est50:*
70: gb_est51:*
71: gb_est52:*
72: gb_est53:*
73: gb_est54:*
74: gb_est55:*
75: gb_est56:*
76: gb_est57:*
77: gb_est58:*
78: gb_est59:*
79: gb_est60:*
80: gb_est61:*
81: gb_est62:*
82: gb_est63:*
83: gb_est64:*
84: gb_est65:*
85: gb_est66:*
86: gb_est67:*
87: gb_est68:*
88: gb_est69:*
89: gb_est70:*
90: gb_est71:*
91: gb_est72:*
92: gb_est73:*
93: gb_est74:*
94: gb_est75:*
95: gb_est76:*
96: gb_est77:*
97: gb_est78:*
98: gb_est79:*
99: gb_est80:*
100: gb_est81:*
101: gb_est82:*
102: gb_est83:*
103: gb_est84:*
104: gb_est85:*
105: gb_est86:*
106: gb_est87:*
107: gb_est88:*
108: gb_est89:*
109: gb_est90:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result

Query

No. Score Match Length DB ID						Description	
C	1	57.4	4.7	925	82	CNS0091P	AL053013 Drosophila
C	2	55.2	4.6	925	82	CNS0091P	AL053013 Drosophila
C	3	51.8	4.3	535	74	AV390505	AV390505 Drosophila
C	4	50.2	4.2	844	82	CNS0052P	AL056652 Drosophila
C	5	49	4.1	910	82	CNS0062P	AL056529 Drosophila
C	6	47	3.9	846	82	CNS010RJ	AL099337 Drosophila
C	7	46.6	3.9	1101	83	CNS0181E	AL108764 Drosophila
C	8	46.4	3.8	935	82	CNS006XK	AL066051 Drosophila
C	9	46.2	3.8	645	82	CNS01213	AL101589 Drosophila
C	10	45.8	3.8	935	82	CNS006XK	AL066051 Drosophila
C	11	45.4	3.8	591	83	AW128786	AW128786 Drosophila
C	12	45.2	3.7	744	83	CNS0172K	AL108658 Drosophila
C	13	45.2	3.7	1101	83	CNS0172X	AL108658 Drosophila
C	14	45	3.7	315	51	AI712257	AI712257 Drosophila
C	15	45	3.7	846	82	CNS0108J	AL099337 Drosophila
C	16	44.8	3.7	932	82	CNS0072Q	AL066742 Drosophila
C	17	44	3.6	843	82	CNS00C51	AL059666 Drosophila
C	18	43.8	3.6	427	51	AI736678	AI736678 Drosophila
C	19	43.8	3.6	512	74	AV396794	AV396794 Drosophila
C	20	43.8	3.6	521	74	AV397116	AV397116 Drosophila
C	21	43.8	3.6	526	74	AV394827	AV394827 Drosophila
C	22	43.6	3.6	1101	83	CNS0160E	AL107216 Drosophila
C	23	43.2	3.6	439	91	AQ084003	AQ084003 Drosophila
C	24	43.2	3.6	798	82	CNS000AJA	AL055851 Drosophila
C	25	43	3.6	413	74	AV396947	AV396947 Drosophila
C	26	41.8	3.5	414	80	AW286084	AW286084 Drosophila
C	27	41.8	3.5	465	80	AW287007	AW287007 Drosophila
C	28	41.8	3.5	530	74	AV396353	AV396353 Drosophila
C	29	41.8	3.5	839	82	CNS00072N	AL056742 Drosophila
C	30	41.6	3.5	932	82	CNS0072N	AL066721 Drosophila
C	31	41.6	3.4	506	74	AW188713	AW188713 Drosophila
C	32	41.6	3.4	870	82	CNS006E2	AL078875 Drosophila
C	33	41.4	3.4	1101	82	CNS000LX	AL103945 Drosophila
C	34	41.4	3.4	1201	83	CNS014BJ	AL181194 Drosophila
C	35	41.4	3.4	639	60	AI812194	AI812194 Drosophila
C	36	41.2	3.4	774	105	AQ0377321	AQ0377321 Drosophila
C	37	41.2	3.4	910	82	CNS0060N	AL066529 Drosophila
C	38	41	3.4	605	46	AI461529	AI461529 Drosophila
C	39	41	3.4	1101	83	CNS0175Y	AL108460 Drosophila
C	40	40.8	3.4	384	64	AW054773	AW054773 Drosophila
C	41	40.8	3.4	385	21	D48746	D48746 Drosophila
C	42	40.8	3.4	437	74	AV395790	AV395790 Drosophila
C	43	40.6	3.4	576	74	AV396405	AV396405 Drosophila
C	44	40.4	3.3	504	91	AQ0848186	AQ0848186 Drosophila
C	45	40.2	3.3	916	83	CNS015SO	AL105860 Drosophila

ALIGNMENTS

Query Match	Best Local Similarity	Matches	4.7% Conservative	Score 57.4; 175;	DB 82; Mismatches 142;	Length 925; Indels 0;	Gaps 0;
442	CCCCGCTGTCGCCGATGAGCGCAGAGTCGGCGGTGACTGCCGCCAGAGTGGCGCTTAC	501	11.9% 43;	Pred. No. 0.058; 175;	Indels 0; 142;	Gaps 0; 142;	
920	SCGCGSCGSCBSCSSSSMTSSNSBSCSSSSBSSTSSMSBSSSSSSSSSGTS	861					
502	CACCTGATGCAACCCAGAGGGGGCCCCGCGTGCATGATGGGGGGGCTCCCGCGCTGAA	561					
860	SACVKNASSSCCGCGCGMABCMCSSSSSSCCGASARCVKVRASGAGKRGSSGGASA	801					
562	CGGCGCGACGTCGATGATCGCGCGCGACCGCGCGCTACACGACGCCGATGCC	621					
800	SHSSSSACBSSSSSCASCMASASSSSSSASRSRSGAGGAGGASSSRSSSSSSASAGS	741					
622	AACGGCATGGCGCGACCGTTACGGCTTACGATCATCAACATGACAAACTTCGGAACTC	681					
740	VVSSASSSSSSSSSVSSSVASMSBSCSSBSSSSASASSSSSSSSSKASCSCCCTSMSC	681					
682	GACGCCGAGCTTCGCGCGGATCCACACGCGTACATCGGCTACGAGCTCGAGGCT	741					
680	SCGTASMSARASSSSSSSSSCSSSMKASSASSSSSSSSSSSSSGGASGCBMSBG	621					

FEATURES
source
Location/Qualifiers
1. 846
/organism="Drosophila melanogaster"
/plasmid="pBelobAC11"
/db_xref="taxon:7227"
/clone_11b="DrosBAC"
/clone="BACN04N13"
/note="end : 17"

BASE COUNT 300 a 47 c 40 g 175 t 284 others

ORIGIN

Query Match 3.9%; Score 47; DB 82; Length 846;
Best Local Similarity 14.0%; Pred. No. 5.8;
Matches 30; Conservative 11; Mismatches 73; Indels 0; Gaps 0;

OY 995 GGCTGACCAACGCGACGATGCTGAGCTGCGACCATGCGCGGCGG 1054
DB 801 SSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 742
OY 1055 CGTGGCGGTGATCGGACGACCAAGCTTTGACGACGAGGCGGTACTGT 1114
DB 741 SASASTSTSTGCGSSBSSTGTSBTTBTSSBSBSSTSSSTSTBTB 682
OY 1115 CCGAAGGCGTGGCCACGACCTGGGGTCCCTACCGACGCCCGACCGCTGGCCT 1174
DB 681 SSSSSSTSTSTSTSTSTSSSTSSSTSSSTSSSTSSSTSSSTSS 622
OY 1175 GACTGCGCGCGCTGCTACGCGACGACGACGACGACGACGACG 1208
DB 621 SSSSCTSSSSSSSSSCTGCGCCSCCCCTC 588

RESULT 7
LOCUS CNS0181E/C
DEFINITION Drosophila melanogaster genome survey sequence SP6 end of BAC
fly, genomic survey sequence.
ACCESSION AL108764
VERSION AL108764.1 GI:5629068
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Séquençage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (BDGP) -
http://www.edgp.edl.ac.uk. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

FEATURES
source
Location/Qualifiers
1. 801
/organism="Drosophila melanogaster"
/plasmid="pBelobAC11"
/db_xref="taxon:7227"
/clone_11b="DrosBAC"
/clone="BACN37H05"
/note="end : SP6"

BASE COUNT 61 a 215 c 301 g 235 t 289 others

ORIGIN

Query Match 3.9%; Score 46.6; DB 83; Length 1101;
Best Local Similarity 40.2%; Pred. No. 7.2;
Matches 101; Conservative 40; Mismatches 109; Indels 1; Gaps 1;

OY 389 CCGGACACGATGATGCTGACGACGACGACGACGACGACGACGACG 448
DB 308 CCGGACGATGATGCTGACGACGACGACGACGACGACGACGACG 249
OY 449 TTGCCCCGATGACGACGACGACGACGACGACGACGACGACGACG 508
DB 248 AAGCGSWSGVGACACAGGCGGAGACCGCCGACGACGACGACGACG 189
OY 509 TCGGACACGACGACGACGACGACGACGACGACGACGACGACGACG 568
DB 188 CCGGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 130
OY 569 ACCTGCTGATGCTGACGACGACGACGACGACGACGACGACGACG 628
DB 129 ACGMAAGGAGAGAGGCGGAGCGGACGCGACGCGCGCGCGCGCG 70
OY 629 TGGCGCGACG 639
DB 69 AGCGCCCAAC 59

RESULT 8
LOCUS CNS006XK
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR14N09 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL066051
VERSION AL066051.1 GI:4945019
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 935)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Séquençage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazuhiro Oseogawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
Location/Qualifiers
1. 935
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_11b="RPCI-98"
/clone="BACR14N09"
/note="end : T7"

BASE COUNT 257 a 170 c 162 g 96 t 250 others

ORIGIN

Query Match 3.8%; Score 46.4; DB 82; Length 935;

[illegible]

```

RESULT 11
Locus      AM128786
Definition  zbrafish washu mpmg est danio rerio cdna 5' similar to
            tr:Q16796 Q16796 NMD(P) TRANSHYDROGENASE PRECURSOR ; mRNA
Accession  AM128786
Version    AM128786
Keywords   zbrafish
Source     zbrafish
Organism   Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
            Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
            Cyprinoidae; Cyprinidae; Rasbora; Danio.
Reference  1 (bases 1 to 591)
            Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M.,
            Eddy,S., Hillier,L., Kucab,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Stepien,M., Theising,R., Allen,M., Bowers,X.,
            Person,B., Smaller,T., Gibbons,M., Pape,D., Harvey,N., Schuck,R.,
            Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterson,D. and Wilson,R.
            Washu zbrafish EST Project 1998
            Unpublished (1998)
            On Jul 8, 1999 this sequence version replaced gi:54232375.
Title      JOURNAL
Comment    Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel.: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            CDNA Library Preparation: Matthew Clark, CDNA Library Arrayed by:
            Matthew Clark, DNA Sequencing by: Washington University Genome
            Sequencing Center Clone distribution: Genome Systems, St. Louis,
            Missouri (web address: www.genomesystems.com) (email contact:
            info@genomesystems.com) and Research Genetics, Huntsville, Alabama
            (web address: www.resgen.com) (email contact: info@resgen.com) and
            ResourceCenter@umprimaridenbank, Berlin, Germany (web address:
            www.rcpd.de)
            Seq primer: T3 RT from Amersham
            High quality sequence stop: 460.
Location/Qualifiers
            1..591

```

```
FEATURES
source      location/qualifiers
1. 591
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone_lib="zebrafish washu MPING EST"
/sex="mixed"
/tissue_type="26 somite embryos, adult livers, shield
```

stage embryos" /lab host="XLI-blue MRF" /note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; 1st strand cDNA was primed with a Not I - Oligo(dT)15 primer [5' pGACTAGTTCAGATCGCGACGCCGCCCTTTTCTTTTCTT3']; double-stranded cDNA was ligated to Sal I adaptors (BRL), digested with Not I and cloned into the Not I and Sal I sites of the pSPORT1 vector (BRL). Library was constructed by Matthew Clark (Lehrnach lab; ICGF, London and Max Planck Institut fuer Molekulare Genetik, Berlin). cDNAs for EST analysis were selected following oligonucleotide hybridization fingerprinting of arrayed clones from zebrafish late somitogenesis (26 ss), adult liver or embryonic shield stage (5.6 h) libraries. Fingerprint data were used to computationally cluster cDNAs, and a single cDNA from each cluster was chosen for sequencing. In some cases multiple members of the same cluster were sequenced to assess clustering parameters or single clones were sequenced additional times to assess quality control."

Query Match 98% Score 45.4; Dn 69; Length 591;
Best Local Similarity 50.2%; Pred No 11;
Matches 112; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

2 TCTTGCAGATTAAATGCAACTTTCTTCACACTGAACCGGTACAAGTAGATGAGAGGGATTATCA 61
Db 185 TCTTAAAGCACATCATCAGCCCTCAACGCCCTCACTAAGTACAGGCACTCTTACCATAAGC 244

OY 62 TGCGGdTGGTAATTCGCGACCGAGACCAAACAGAAATTCGGGTGCCATCACCCGG 121
Db 245 TGACGGTGtGGGGTTCTCTAAGAAMATCTTTCAGAAATGAGACGTCGAGTGGCCATCTCTCCGT 304

OY 122 CCGGGGTGCGGAAATCAACCCGTCGTGGSCANTAGACTCTCATCCAGGCAGATGCCGAG 181
Db 305 CCGGATGTGAAGACTTCTATCAAGCAGGGCTTTAAATGTGTGTTGGAGTCTCGCGGGAG 364

OY 182 AGGGCTCGGCTATCACCGACCGGATTTCAAGCGCGCAGCGGC 224
Db 365 AATGTCCAAGTTCTCTGATGACATGTACACAAAAGCAGAGAC 407

RESULT 12
CNBS017ZK 744 bp DNA GSS 26-JUL-1999
LOCUS CNBS017ZK
DEFINITION Drosophila melanogaster genome survey sequence Sp6 end of BAC
BAON37B10 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION ALI08698
VERSION ALI08698.1 GI:5629002
KEYWORDS GSS.

SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 744)
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
Web : www.genoscope.cns.fr)

COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billard at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pae JBAC11.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2000, 14:57:02 ; Search time 8627.09 Seconds

(without alignments)
-139.371 Million cell updates/sec

Title: US-09-362-485-10
Perfect score: 1236
Sequence: 1 ATCTTGAGATTATCGAAC.....GGAGAGCATGATGCGGCC 1236

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenBml:*

- 1: gb_ba1:*
- 2: gb_ba2:*
- 3: gb_om:*
- 4: gb_ov:*
- 5: gb_pat:*
- 6: gb_ph:*
- 7: gb_pl:*
- 8: gb_pl2:*
- 9: gb_pr1:*
- 10: gb_pr2:*
- 11: gb_pr3:*
- 12: gb_ro:*
- 13: gb_sts:*
- 14: gb_sy:*
- 15: gb_un:*
- 16: gb_vl:*
- 17: em_fun:*
- 18: em_hum1:*
- 19: em_hum2:*
- 20: em_in:*
- 21: em_om:*
- 22: em_or:*
- 23: em_ov:*
- 24: em_pat:*
- 25: em_ph:*
- 26: em_pl:*
- 27: em_ro:*
- 28: em_sts:*
- 29: em_sy:*
- 30: em_un:*
- 31: em_vl:*
- 32: gb_hc91:*
- 33: gb_hc92:*
- 34: gb_in1:*
- 35: gb_in2:*
- 36: em_ba1:*
- 37: em_ba2:*
- 38: em_hum3:*
- 39: em_hum4:*
- 40: gb_pr4:*
- 41: gb_hc93:*
- 42: gb_hc94:*
- 43: gb_hc95:*
- 44: gb_hc96:*

45: gb_hc97:*

46: em_hc91:*

47: em_hc92:*

48: em_hc93:*

49: em_hums:*

50: gb_pl3:*

51: gb_pr5:*

52: gb_hc98:*

53: gb_hc99:*

54: gb_hc10:*

55: gb_hc11:*

56: gb_hc12:*

57: gb_hc13:*

58: gb_hc14:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1236	100.0	1236	5	A87613	A87613 Sequence 10
2	1236	100.0	1236	5	A89753	A89753 Sequence 10
3	1234.4	99.9	1237	5	A87607	A87607 Sequence 4
4	1234.4	99.9	1237	5	A89747	A89747 Sequence 4
5	1234.4	99.9	56414	1	MTV002	AL008967 Mycobacte
6	1232	99.7	1235	5	A87606	A87606 Sequence 3
7	1232	99.7	1235	5	A89746	A89746 Sequence 3
8	1226.4	99.2	1228	5	A87608	A87608 Sequence 5
9	1226.4	99.2	1228	5	A89748	A89748 Sequence 5
10	1223	98.9	1235	5	A87609	A87609 Sequence 6
11	1223	98.9	1235	5	A89749	A89749 Sequence 6
12	1222.4	98.9	1235	5	A87611	A87611 Sequence 8
13	1222.4	98.9	1235	5	A89751	A89751 Sequence 8
14	1220	98.7	1245	5	A87605	A87605 Sequence 2
15	1220	98.7	1245	5	A89745	A89745 Sequence 2
16	1220	98.7	1260	5	A87604	A87604 Sequence 1
17	1220	98.7	1260	5	A89744	A89744 Sequence 1
18	1220	98.7	2584	1	MTALADH	X63069 M.tuberculo
19	1215.4	98.3	1229	5	A87610	A87610 Sequence 7
20	1215.4	98.3	1229	5	A89750	A89750 Sequence 7
21	1208	97.7	1209	5	A87612	A87612 Sequence 9
22	1201	97.2	1208	5	A89752	A89752 Sequence 9
23	1194	96.6	1194	2	MT092472	U92472 Mycobacteri
24	698	56.5	720	5	A89759	A89759 Sequence 16
25	538.4	43.6	40745	1	SC151	AF109848 Streptomy
26	467.6	37.8	682	5	A89758	A89758 Sequence 15
27	467.6	37.8	682	5	A89752	A89752 Sequence 19
28	377	30.5	8046	2	AF049107	AF049107 Myxococcu
29	351.2	28.4	11514	2	AE002028	AE002028 Delnococt
30	331.4	26.8	2780	1	RLK238118	AJ238118 Rhizobium
31	330.8	26.8	1585	2	AF070716	AF070716 Vibrio pr
32	312.8	25.3	1736	1	BACALDHC	M33299 B.stearothe
33	293	23.7	1477	2	AF070715	AF070715 Shewanell
34	288.4	23.3	2113	1	D37807	D37807 Phormidium
35	277.2	22.4	1125	5	E01847	E01847 DNA encodin
36	258	20.9	137740	1	D90900	D90900 Synchocyst
37	243.2	19.7	1375	1	AB013821	AB013821 Enterobac
38	243	19.7	1776	1	BACALDH	M33298 B.sphaericu
39	235.6	19.1	1709	1	BACALD	L20916 Bacillus su
40	235.6	19.1	12316	1	BSR82015	282015 B.subtilis
41	235.6	19.1	217420	1	BSR80017	299120 Bacillus su
42	218.2	17.7	1714	2	AF070714	AF070714 Carnobact
43	107.2	8.7	10512	2	AE001565	AE001565 Helicobac
44	99.2	8.0	10512	2	AE000639	AE000639 Helicobac
45	74.2	6.0	3160	1	RR05294	U05294 Rhodospirill

ALIGNMENTS

RESULT 1
 LOCUS AB7613 1236 bp DNA PAT 22-JAN-2000
 DEFINITION Sequence 10 from Patent WO9836089.
 ACCESSION AB7613
 VERSION AB7613.1 GI:6736253
 KEYWORDS
 SOURCE
 ORGANISM
 unclassified.
 unclassified.
 unclassified.
 REFERENCE
 1 (bases 1 to 1236)
 Flohe, L. and Singh, M.
 TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
 Patent: WO 9836089-A 20-AUG-1998;
 JOURNAL
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source
 1..1236
 Location/Qualifiers
 BASE COUNT 236 a 395 c 385 g 220 t
 ORIGIN

Query Match 100.0%; Score 1236; DB 5; Length 1236;
 Best Local Similarity 100.0%; Pred. No. 9.4e-170;
 Matches 1236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATCTTGCAATTAATGACATTTCTTACACAGCTGACAGTATCGAGAGGGTAATC 60
 1 ATCTTGCAATTAATGACATTTCTTACACAGCTGACAGTATCGAGAGGGTAATC 60
 Db 1 ATCTTGCAATTAATGACATTTCTTACACAGCTGACAGTATCGAGAGGGTAATC 60
 QY 61 ATGCGGTGGTATTCGACGACGACCAAAACAAAGAAATCCGGGTGCATCAACCCG 120
 61 ATGCGGTGGTATTCGACGACGACCAAAACAAAGAAATCCGGGTGCATCAACCCG 120
 Db 61 ATGCGGTGGTATTCGACGACGACCAAAACAAAGAAATCCGGGTGCATCAACCCG 120
 QY 121 GCGCGGTGGCGAATAACCCGCTGCGCAATGAGTGCATCATCAGCAGTGCCTGA 180
 121 GCGCGGTGGCGAATAACCCGCTGCGCAATGAGTGCATCATCAGCAGTGCCTGA 180
 Db 121 GCGCGGTGGCGAATAACCCGCTGCGCAATGAGTGCATCATCAGCAGTGCCTGA 180
 QY 181 GAGGCTCGGTATACACGACGCGATTTCAAGGCGGACGCGCAATGCTCGGCACC 240
 181 GAGGCTCGGTATACACGACGCGATTTCAAGGCGGACGCGCAATGCTCGGCACC 240
 Db 181 GAGGCTCGGTATACACGACGCGATTTCAAGGCGGACGCGCAATGCTCGGCACC 240
 QY 241 GCGGACAGGTGTGGCGCAGCGTGTATTTATGCTCAAGTCAAAAGAACCGATAGCGCG 300
 241 GCGGACAGGTGTGGCGCAGCGTGTATTTATGCTCAAGTCAAAAGAACCGATAGCGCG 300
 Db 241 GCGGACAGGTGTGGCGCAGCGTGTATTTATGCTCAAGTCAAAAGAACCGATAGCGCG 300
 QY 301 GAATAGCGCGCTGCGAAGCGGAGATCTGTACGTTCTTGATTTGGCCCGGTCA 360
 301 GAATAGCGCGCTGCGAAGCGGAGATCTGTACGTTCTTGATTTGGCCCGGTCA 360
 Db 301 GAATAGCGCGCTGCGAAGCGGAGATCTGTACGTTCTTGATTTGGCCCGGTCA 360
 QY 361 CGTGTGACGATCGTGTGTGATTCGCGCAGCAGCTCAATTCCTCTACAGACCGCTC 420
 361 CGTGTGACGATCGTGTGTGATTCGCGCAGCAGCTCAATTCCTCTACAGACCGCTC 420
 Db 361 CGTGTGACGATCGTGTGTGATTCGCGCAGCAGCTCAATTCCTCTACAGACCGCTC 420
 QY 421 CAGACCGCGAGCGGCGCTACCTCTGTTGCCCCGATGAGCAAGTCCCGTGCATTC 480
 421 CAGACCGCGAGCGGCGCTACCTCTGTTGCCCCGATGAGCAAGTCCCGTGCATTC 480
 Db 421 CAGACCGCGAGCGGCGCTACCTCTGTTGCCCCGATGAGCAAGTCCCGTGCATTC 480
 QY 481 GCGCGCGAGTTGGGCTTACCAACGATGAGCAAGCGGCGCGGTGTGCGTATG 540
 481 GCGCGCGAGTTGGGCTTACCAACGATGAGCAAGCGGCGCGGTGTGCGTATG 540
 Db 481 GCGCGCGAGTTGGGCTTACCAACGATGAGCAAGCGGCGCGGTGTGCGTATG 540
 QY 541 GCGCGGTGGCGGCTGCAACCGCGGAGCTGTGTATCGGCGCGGCAACCGCGCGC 600
 541 GCGCGGTGGCGGCTGCAACCGCGGAGCTGTGTATCGGCGCGGCAACCGCGCGC 600
 Db 541 GCGCGGTGGCGGCTGCAACCGCGGAGCTGTGTATCGGCGCGGCAACCGCGCGC 600
 QY 601 TACAGCAGACCCGATCGCAACGGGATGCGCGACCGTATACGGTTTAAACATCAAC 660
 601 TACAGCAGACCCGATCGCAACGGGATGCGCGACCGTATACGGTTTAAACATCAAC 660
 Db 601 TACAGCAGACCCGATCGCAACGGGATGCGCGACCGTATACGGTTTAAACATCAAC 660
 QY 661 ATGGAACAACCTTGGCACTCGACGCGGAGTCTGGCGCGGATCAACACTCGCTACTCA 720

Db 661 ATGGAACAACCTTGGCACTCGACGCGGAGTCTGGCGCGGATCAACACTCGCTACTCA 720
 QY 721 TCGGCTACAGAGCTCGAGGGTGGCGTCAACGTCAGCGCACTGATGATGGGCGCTCCG 780
 721 TCGGCTACAGAGCTCGAGGGTGGCGTCAACGTCAGCGCACTGATGATGGGCGCTCCG 780
 Db 721 TCGGCTACAGAGCTCGAGGGTGGCGTCAACGTCAGCGCACTGATGATGGGCGCTCCG 780
 QY 781 GTGCGAGCGCGCAAGGACCCAAATATAGTCTCGAATTCACCTTGTGCGCATATGAACCA 840
 781 GTGCGAGCGCGCAAGGACCCAAATATAGTCTCGAATTCACCTTGTGCGCATATGAACCA 840
 Db 781 GTGCGAGCGCGCAAGGACCCAAATATAGTCTCGAATTCACCTTGTGCGCATATGAACCA 840
 QY 841 GGTGCGGTACTGTGTATATATGATGATGATGATGATGATGATGATGATGATGATGATG 900
 841 GGTGCGGTACTGTGTATATATGATGATGATGATGATGATGATGATGATGATGATGATG 900
 Db 841 GGTGCGGTACTGTGTATATATGATGATGATGATGATGATGATGATGATGATGATGATG 900
 QY 901 ACCACCTACGACACACCGGAGCTGCGGTGCGACAGACAGCGTGTCTTACTGCGTGGGAAC 960
 901 ACCACCTACGACACACCGGAGCTGCGGTGCGACAGACAGCGTGTCTTACTGCGTGGGAAC 960
 Db 901 ACCACCTACGACACACCGGAGCTGCGGTGCGACAGACAGCGTGTCTTACTGCGTGGGAAC 960
 QY 961 ATGCGCGCTGCTGCTGCGAAGACGTGACCTACGCGCTGACCAACGCGAGATGCGGTAT 1020
 961 ATGCGCGCTGCTGCTGCGAAGACGTGACCTACGCGCTGACCAACGCGAGATGCGGTAT 1020
 Db 961 ATGCGCGCTGCTGCTGCGAAGACGTGACCTACGCGCTGACCAACGCGAGATGCGGTAT 1020
 QY 1021 GTGCTCGAGCTTGGCGAAGCATGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
 1021 GTGCTCGAGCTTGGCGAAGCATGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
 Db 1021 GTGCTCGAGCTTGGCGAAGCATGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
 QY 1081 AAAGGTCTTTGACGACGAGGAGGCGCTTACTGTCCGAGAGGAGTGGCGACGACCTGGGG 1140
 1081 AAAGGTCTTTGACGACGAGGAGGCGCTTACTGTCCGAGAGGAGTGGCGACGACCTGGGG 1140
 Db 1081 AAAGGTCTTTGACGACGAGGAGGCGCTTACTGTCCGAGAGGAGTGGCGACGACCTGGGG 1140
 QY 1141 GTGCGGTTCACCGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
 1141 GTGCGGTTCACCGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
 Db 1141 GTGCGGTTCACCGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
 QY 1201 CACAGTGGGAGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1236
 1201 CACAGTGGGAGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1236
 Db 1201 CACAGTGGGAGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1236

RESULT 2
 LOCUS AB9753 1236 bp DNA PAT 22-JAN-2000
 DEFINITION Sequence 10 from Patent WO9832862.
 ACCESSION AB9753
 VERSION AB9753.1 GI:6738287
 KEYWORDS
 SOURCE
 ORGANISM
 unclassified.
 unclassified.
 unclassified.
 REFERENCE
 1 (bases 1 to 1236)
 Flohe, L. and Singh, M.
 L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
 Patent: WO 9832862-A 30-JUL-1998;
 JOURNAL
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source
 1..1236
 Location/Qualifiers
 BASE COUNT 236 a 395 c 385 g 220 t
 ORIGIN

Query Match 100.0%; Score 1236; DB 5; Length 1236;
 Best Local Similarity 100.0%; Pred. No. 9.4e-170;
 Matches 1236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATCTTGCAATTAATGACATTTCTTACACAGCTGACAGTATCGAGAGGGTAATC 60
 1 ATCTTGCAATTAATGACATTTCTTACACAGCTGACAGTATCGAGAGGGTAATC 60
 Db 1 ATCTTGCAATTAATGACATTTCTTACACAGCTGACAGTATCGAGAGGGTAATC 60
 QY 61 ATGCGGTGGTATTCGACGACGACCAAAACAAAGAAATCCGGGTGCATCAACCCG 120
 61 ATGCGGTGGTATTCGACGACGACCAAAACAAAGAAATCCGGGTGCATCAACCCG 120
 Db 61 ATGCGGTGGTATTCGACGACGACCAAAACAAAGAAATCCGGGTGCATCAACCCG 120

```

Db 61 ATGCGGCTGATTCCGACGAGACCAAAACAAACAAATTCGGGTGGCATACACCCG 120
QY 121 GCGGGGTGGCGAATATACCCGTCGTGGCATAGAGTCTCATTCAGGACAGTGGCGA 180
Db 121 GCGGGGTGGCGAATATACCCGTCGTGGCATAGAGTCTCATTCAGGACAGTGGCGA 180
QY 181 GAGGGCTGCTATACGACGAGGATTCAAAGGCGGCGGCAATGGTCGGACCC 240
Db 181 GAGGGCTGCTATACGACGAGGATTCAAAGGCGGCGGCAATGGTCGGACCC 240
QY 241 GCCGACCAAGTGTGGCGGACGCTGATTTATGCTCAAGTCAAAAGACCAATAGCGCG 300
Db 241 GCCGACCAAGTGTGGCGGACGCTGATTTATGCTCAAGTCAAAAGACCAATAGCGCG 300
QY 301 GAATAGCGGCGCTGGGACACGGGAGATCTGTACGTTCTTGATTTGGCCGGGTGA 360
Db 301 GAATAGCGGCGCTGGGACACGGGAGATCTGTACGTTCTTGATTTGGCCGGGTGA 360
QY 361 CGTGTGACACGATCGTGTGATTCGGGACGACGTCATTAATGCTACGACGACGTC 420
Db 361 CGTGTGACACGATCGTGTGATTCGGGACGACGTCATTAATGCTACGACGACGTC 420
QY 421 CAGACCGCCGACGGGCGACTACCCCTGCTGCCCCGATAGCGAGTGGCCGGTGCATC 480
Db 421 CAGACCGCCGACGGGCGACTACCCCTGCTGCCCCGATAGCGAGTGGCCGGTGCATC 480
QY 481 GCGGCCAAGTGGGCGCTTACCACTGATGCGAACGCCAGGGGCGCGGTGTGCTGATG 540
Db 481 GCGGCCAAGTGGGCGCTTACCACTGATGCGAACGCCAGGGGCGCGGTGTGCTGATG 540
QY 541 GCGGGGTCGCGGCGCTTCGACACGGGCGGACGTCGTGATGCGGCGCGGACCGCGCG 600
Db 541 GCGGGGTCGCGGCGCTTCGACACGGGCGGACGTCGTGATGCGGCGCGGACCGCGCG 600
QY 601 TACACGACGACCGCGCATCGCCAAAGGCGCATGGGCGGACCGTGAAGTCTAGACATCA 660
Db 601 TACACGACGACCGCGCATCGCCAAAGGCGCATGGGCGGACCGTGAAGTCTAGACATCA 660
QY 661 ATCGAACAACTTCGGAATCTCGACGCGGAGTTCTGGGCGGATTCACACTGCTACTCA 720
Db 661 ATCGAACAACTTCGGAATCTCGACGCGGAGTTCTGGGCGGATTCACACTGCTACTCA 720
QY 721 TCGGCTTACGAGCTCGAGGTCGCTCAAAAGTGGCGGACGTCGTGATTTGGGCGGCTG 780
Db 721 TCGGCTTACGAGCTCGAGGTCGCTCAAAAGTGGCGGACGTCGTGATTTGGGCGGCTG 780
QY 781 GTGCGACGCGGCAAGGCGACCAATTAATGCTCGAATTCATCTGCGCATATGAACA 840
Db 781 GTGCGACGCGGCAAGGCGACCAATTAATGCTCGAATTCATCTGCGCATATGAACA 840
QY 841 GGTGGGCTACTGTTGATATATGATTCGACACGAGGGGCGTGTTCGAAGGCTACACCG 900
Db 841 GGTGGGCTACTGTTGATATATGATTCGACACGAGGGGCGTGTTCGAAGGCTACACCG 900
QY 901 ACCACCTACGACACCGCGAGTTCGCGGTGACGACGACGCTGTTTACTGCGTGGGAAC 960
Db 901 ACCACCTACGACACCGCGAGTTCGCGGTGACGACGACGCTGTTTACTGCGTGGGAAC 960
QY 961 ATGCGCGGCTCGGTGCGGAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1020
Db 961 ATGCGCGGCTCGGTGCGGAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1020
QY 1021 GTGCTGAGACGTCGACGATGCTGGGCGGCGGTGCGGCTGCAATCGGCGACTAGCC 1080
Db 1021 GTGCTGAGACGTCGACGATGCTGGGCGGCGGTGCGGCTGCAATCGGCGACTAGCC 1080
QY 1081 AAAGGCTTTTCAGACGACGAAAGGGGCTTACTCTCGGAAGGGTGGCCACGACCTGGGG 1140
Db 1081 AAAGGCTTTTCAGACGACGAAAGGGGCTTACTCTCGGAAGGGTGGCCACGACCTGGGG 1140
QY 1141 GTGCGCTTACCGAGCGCGGACGCTGCTGACTCTGCGCGCTGCTTACGCGCGAG 1200
Db 1141 GTGCGCTTACCGAGCGCGGACGCTGCTGACTCTGCGCGCTGCTTACGCGCGAG 1200

```

```

QY 1201 CACAGCTGGGAGTAAGGAAACGATGATGTCGCC 1236
Db 1201 CACAGCTGGGAGTAAGGAAACGATGATGTCGCC 1236

RESULT 3
A87607
LOCUS A87607 1237 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 4 from Patent WO9836089.
ACCESSION A87607
VERSION A87607.1 GI:6736247
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1237)
AUTHORS Flohe, L. and Singh, M.
TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL JOURNAL
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
source
1..1237
Location/Qualifiers
BASE COUNT 236 a 394 c 386 g 221 t
ORIGIN

Query Match 99.9% Score 1234.4; DB 5: Length 1237;
Best Local Similarity 99.9%; Pred. No. 1.6e-169;
Matches 1235; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATCTTGACGATTAAATGAACTTCTTCACTGAGAGCTACAGTATCGAGAGGGTAATC 60
Db 1 ATCTTGACGATTAAATGAACTTCTTCACTGAGAGCTACAGTATCGAGAGGGTAATC 60
QY 61 ATGCGCGTCGTTATTCGCGACGACCAAAACAAAGCAATTCGGGTGGCCATCACCCG 120
Db 61 ATGCGCGTCGTTATTCGCGACGACCAAAACAAAGCAATTCGGGTGGCCATCACCCG 120
QY 121 GCGGCGTCGCGGAACTAACCCGTCGTGGCCATGAGGTGCTATCCAGGAGGTGCGGGA 180
Db 121 GCGGCGTCGCGGAACTAACCCGTCGTGGCCATGAGGTGCTATCCAGGAGGTGCGGGA 180
QY 181 GAGGGCTCGGCTATACCGACGCGGATTCAGAGGCGGACGCGCACTGTCGGCAC 240
Db 181 GAGGGCTCGGCTATACCGACGCGGATTCAGAGGCGGACGCGCACTGTCGGCAC 240
QY 241 GCCGACAGGTCGCGGCGACGCTGATTTATGCTCAAGTCAAAAGCCGATAGCGGCG 300
Db 241 GCCGACAGGTCGCGGCGACGCTGATTTATGCTCAAGTCAAAAGCCGATAGCGGCG 300
QY 301 GAATAGCGGCGCTGCGACACGGGAGATCTGTACGTTCTTGATTTGGCCGGGTGA 360
Db 301 GAATAGCGGCGCTGCGACACGGGAGATCTGTACGTTCTTGATTTGGCCGGGTGA 360
QY 361 CGTGTGACACGATCGTGTGATTCGGGACGACGTCATTAATGCTACGACGACGTC 420
Db 361 CGTGTGACACGATCGTGTGATTCGGGACGACGTCATTAATGCTACGACGACGTC 420
QY 421 CAGACCGCCGACGGGCGACTACCCCTGCTGCCCCGATAGAGGAGTGGCGGTGACATC 480
Db 421 CAGACCGCCGACGGGCGACTACCCCTGCTGCCCCGATAGAGGAGTGGCGGTGACATC 480
QY 481 GCGGCCAAGTGGGCGCTTACCACTGATGCGAACCCAAAGGGGCGCGGTGTGCTGATG 540
Db 481 GCGGCCAAGTGGGCGCTTACCACTGATGCGAACCCAAAGGGGCGCGGTGTGCTGATG 540
QY 541 GCGGGGTCGCGGCGCTTCGACACGGGCGGACGTCGTGATGCGGCGGACCGCGCG 600
Db 541 GCGGGGTCGCGGCGCTTCGACACGGGCGGACGTCGTGATGCGGCGGACCGCGCG 600

```

```

QY 601 TACAAAGCAGCCCGCATGCGCAACGCGATGGGCGCGCGTACGTTACGTTCTAGACATCAAC 660
DB 601 TACAAAGCAGCCCGCATGCGCAACGCGATGGGCGCGCGTACGTTACGTTCTAGACATCAAC 660
QY 661 ATGAGCAAACTTGGCACTCGACGCGGAGTTCTGCGCGGATCCACACTGCTACTCA 720
DB 661 ATGAGCAAACTTGGCACTCGACGCGGAGTTCTGCGCGGATCCACACTGCTACTCA 720
QY 721 TGGGCTAGAGCTGAGGAGTGGCGGCAACGCGGATGGTATGTTGGGCGCGCTG 780
DB 721 TGGGCTAGAGCTGAGGAGTGGCGGCAACGCGGATGGTATGTTGGGCGCGCTG 780
QY 781 GTGCGAGCGCCCAAGGCAACCAATTAGTCTGCAATTCACTTGTGCGCATATGAACA 840
DB 781 GTGCGAGCGCCCAAGGCAACCAATTAGTCTGCAATTCACTTGTGCGCATATGAACA 840
QY 841 GGTGGGCTACTGTTGATATATGCAATCGACAGGCGGCTGTTTGAAGGCTCAGACCG 900
DB 841 GGTGGGCTACTGTTGATATATGCAATCGACAGGCGGCTGTTTGAAGGCTCAGACCG 900
QY 901 ACCACCTAGACGACCGCGAGTTCGCGGTGACAGACGCGTGTCTTACTGCGTGGCAAC 960
DB 901 ACCACCTAGACGACCGCGAGTTCGCGGTGACAGACGCGTGTCTTACTGCGTGGCAAC 960
QY 961 ATGCGCGCTCGGTGCGGAGAGAGAGTGCAGCTACGCGCTACCAACGAGATGCCGTAT 1020
DB 961 ATGCGCGCTCGGTGCGGAGAGAGAGTGCAGCTACGCGCTACCAACGAGATGCCGTAT 1020
QY 1021 GTGCTGACGCTTGGCGAGATGGCTGGGCGGCGGTGCGGATCGGAGTACGCGTACGCG 1080
DB 1021 GTGCTGACGCTTGGCGAGATGGCTGGGCGGCGGTGCGGATCGGAGTACGCGTACGCG 1080
QY 1081 AAGGCTTTGCAAGCAGCAAGGCGGCTTACTGTCCGAAGGCGGTGCGCAACGACTGGG 1140
DB 1081 AAGGCTTTGCAAGCAGCAAGGCGGCTTACTGTCCGAAGGCGGTGCGCAACGACTGGG 1140
QY 1141 GTGCGCTTACGAGCGCGCGGAGCGTGTGCTGACTGTGCGCGCTGTTACGCGGAG 1200
DB 1141 GTGCGCTTACGAGCGCGCGGAGCGTGTGCTGACTGTGCGCGCTGTTACGCGGAG 1200
QY 1201 CACAGCTCGGAGTAAAGGAGGATGATGTGGCC 1236
DB 1201 CACAGCTCGGAGTAAAGGAGGATGATGTGGCC 1236

RESULT 4
LOCUS A89747 1237 bp DNA
DEFINITION Sequence 4 from Patent WO9832862.
ACCESSION A89747
VERSION A89747.1 GI:6738281
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 1237)
AUTHORS Flohe, L. and Singh, M.
TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
JOURNAL Patent: WU 9832862-A 30-JUL-1998;
FLOHE LEOPOLD (DE); SINGH MAHADEVIR (DE)
FEATURES
source
1..1237
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 236 a 394 c 386 g 221 t
ORIGIN
Query Match 99.9%; Score 1234.4; DB 5; Length 1237;
Best Local Similarity 99.9%; Pred. No. 1.6e-169;
Matches 1235; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATCTTGACATTATGAACTTTCTTACACGCTACGAGCGTACGATCGAGAGGGTAATC 60

```

```

DB 1 ATCTTGACATTATGAACTTTCTTACGAGCTACGATATGAGAGGGTATC 60
QY 61 ATGCGCGTGGTATTCGACGAGCAACCAAAACAGAAATTCGGGTGGCATACCCCG 120
DB 61 ATGCGCGTGGTATTCGACGAGCAACCAAAACAGAAATTCGGGTGGCATACCCCG 120
QY 121 GCGGCGTGGGGAATTAACCGTGTGGCATGAGTGTGATCATCAGGAGTGGCGGA 180
DB 121 GCGGCGTGGGGAATTAACCGTGTGGCATGAGTGTGATCATCAGGAGTGGCGGA 180
QY 181 GAGGCGTGGGCTATCAGCAGCGCGGATTTCAAGGCGGAGCGCGCACTGTGGCAC 240
DB 181 GAGGCGTGGGCTATCAGCAGCGCGGATTTCAAGGCGGAGCGCGCACTGTGGCAC 240
QY 241 GCGGACAGAGTGTGGGCGG TGTGATTTATTTGCTCAAGGCTCAAGAAACGATAGCGCG 300
DB 241 GCGGACAGAGTGTGGGCGGCTGATTTATTTGCTCAAGGCTCAAGAAACGATAGCGCG 300
QY 301 GAATAGGCGCGGCTGCGACAGCGGAGATCTTGTTCAGGTTCTTGCATTTGGCGGTCA 360
DB 301 GAATAGGCGCGGCTGCGACAGCGGAGATCTTGTTCAGGTTCTTGCATTTGGCGGTCA 360
QY 361 GGTGCTTGACCGATGCGTTGTTGGATTCGCGCACACGTCATTTGCTATAGACCGTC 420
DB 361 GGTGCTTGACCGATGCGTTGTTGGATTCGCGCACACGTCATTTGCTATAGACCGTC 420
QY 421 CAGACGCGGAGAGGCGGCACTACCCCTGCTTGGCCCGATGAGAGGAATGGCGGTGAC 480
DB 421 CAGACGCGGAGAGGCGGCACTACCCCTGCTTGGCCCGATGAGAGGAATGGCGGTGAC 480
QY 481 GCGGCGGAGGTTGGCTTACCACTGATGCGAACCCAAAGGAGGCGGCGGTGCTGATG 540
DB 481 GCGGCGGAGGTTGGCTTACCACTGATGCGAACCCAAAGGAGGCGGCGGTGCTGATG 540
QY 541 GCGGCGGTTGCCGCGGTGGAACCGGCGGAGTGTGATGCGCGCGGCGGACCGCGG 600
DB 541 GCGGCGGTTGCCGCGGTGGAACCGGCGGAGTGTGATGCGCGCGGCGGACCGCGG 600
QY 601 TACAGGAGCGCGGATGCGCAAGGCGATGGGCGGAGCGTTAGCGTCTAGACATCAAC 660
DB 601 TACAGGAGCGCGGATGCGCAAGGCGATGGGCGGAGCGTTAGCGTCTAGACATCAAC 660
QY 661 ATGCAAAACTTGGCACTGAGCGCGGAGTCTGCGCGGATCCACACTGCTACTCA 720
DB 661 ATGCAAAACTTGGCACTGAGCGCGGAGTCTGCGCGGATCCACACTGCTACTCA 720
QY 721 TCGGCTTACGAGCTGAGAGTCTGTCAAAGCTGCGACCTGTGATGGGCGGCTCTG 780
DB 721 TCGGCTTACGAGCTGAGAGTCTGTCAAAGCTGCGACCTGTGATGGGCGGCTCTG 780
QY 781 GTGCGAGGCGCGCAAGGCAACCAATTAGTCTGCAATTCACTTGTGCGCATATGAACA 840
DB 781 GTGCGAGGCGCGCAAGGCAACCAATTAGTCTGCAATTCACTTGTGCGCATATGAACA 840
QY 841 GGTGGGCTACTGTTGATATATGCAATCGACAGGCGGCTGTTTGAAGGCTCAGACCG 900
DB 841 GGTGGGCTACTGTTGATATATGCAATCGACAGGCGGCTGTTTGAAGGCTCAGACCG 900
QY 901 ACCACCTAGACGACCGCGAGTTCGCGGTGACAGACGCGTGTCTTACTGCGTGGCAAC 960
DB 901 ACCACCTAGACGACCGCGAGTTCGCGGTGACAGACGCGTGTCTTACTGCGTGGCAAC 960
QY 961 ATGCGCGCTCGGTGCGGAGAGAGTGCAGCTACGCGCTACCAACGAGATGCCGTAT 1020
DB 961 ATGCGCGCTCGGTGCGGAGAGAGTGCAGCTACGCGCTACCAACGAGATGCCGTAT 1020
QY 1021 GTGCTGACGCTTGGCGAGATGGCTGGGCGGCGGTGCGGATCGGAGTACGCGTACG 1080
DB 1021 GTGCTGACGCTTGGCGAGATGGCTGGGCGGCGGTGCGGATCGGAGTACGCGTACG 1080
QY 1081 AAGGCTTTGCAAGCAGCAAGGCGGCTTACTGTCCGAAGGCGGTGCGACCGACTGGGG 1140
DB 1081 AAGGCTTTGCAAGCAGCAAGGCGGCTTACTGTCCGAAGGCGGTGCGACCGACTGGGG 1140

```

Db	Accession	Source	Organism	Keywords	Version	Accession	Definition	Locust	Result
Db	1081	AAAGTCTTTTCGACGACAGGAGGCGCTTACTGTGCCAAAGGGATGGCCACCGACTGTGGG	1140						
Qy	1141	GTGGCGTTACACGAGCCCGGACGCGTGTGGCCCTGACTCTGGGCGCCCTGTACCGGAG	1200						
Db	1141	GTGGCGTTACACGAGCCCGGACGCGTGTGGCCCTGACTCTGGGCGCCCTGTACCGGAG	1200						
Qy	1201	CACAGCTGGAGTAAAGGAGGAGCGATGATGTGCGCC	1236						
Db	1201	CACAGCTGGAGTAAAGGAGGAGCGATGATGTGCGCC	1236						
Result	5								
LOCUS	MTV002	56414 bp	DNA	BC1					
DEFINITION	Mycobacterium tuberculosis H37Rv complete genome; segment 122/162.								
ACCESSION	AL008867	AL123456							
VERSION	AL008867.1	GI:3261491							
KEYWORDS	Mycobacterium tuberculosis.								
SOURCE	Mycobacterium tuberculosis								
ORGANISM	Bacteria: Firmicutes: Actinobacteria: Actinobacteridae: Bacteria: Firmicutes: Actinobacteriineae; Mycobacteriaceae;								
REFERENCE	1 (bases 1 to 56414)								
AUTHORS	Cole,S.T., Brosch,R., Parkhill,J., Garnier,T., Churcher,C., Harris,D., Gordon,S.V., Eigmeier,K., Gas,S., Barry III,C.E., Teale,F., Badcock,K., Basham,D., Brown,D., Chillingworth,T., Connor,R., Davies,R., Devlin,K., Feltwell,T., Gentles,S., Hamlin,N., Holroyd,S., Hornsby,T., Jagels,K., Krogh,A., McLean,J., Moule,S., Murphy,L., Oliver,S., Osborne,J., Quail,M.A., Randle,M., Rasmussen,M., Rogers,J., Rutter,S., Seeger,K., Skelton,S., Squares,S., Squares,R., Sultson,J.E., Taylor,K., Whitehead,S. and Barrett,B.G.								
TITLE	Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence								
JOURNAL	Nature 393 (6685), 537-544 (1998)								
MEDLINE	98295987								
REMARK	Erratum:[published erratum appears in Nature 1998 Nov 12;396(6707):1901]								
REFERENCE	2 (bases 1 to 56414)								
AUTHORS	Parkhill,J.								
TITLE	Direct Submission								
JOURNAL	Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux, 75712 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk								
COMMENT	On Jun 27, 1998 this sequence version replaced gi:2624256.								
Notes:	Details of M. tuberculosis sequencing at the Sanger Centre are available on the World Wide Web. (URL, http://www.sanger.ac.uk/Projects/M_tuberculosis/) CDS have been renumbered from the original cosmid submissions but the old gene designations are in brackets after the new gene numbers. Gene prediction was based on a Hidden Markov Model of TB genes implemented in TBparse (Krogh) supplemented with visual inspection of positional base preference in codons, especially where there is an increase in the observed/expected third position G + C. CAUTION: In some cases we may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, or ttg) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.								
FEATURES	Location/Qualifiers								
SOURCE	1..56414								
	/organism="Mycobacterium tuberculosis"								
	/strain="H37Rv"								
	/db_xref="taxon:1773"								
	<1..>56414								
	/note="fragment designated v002. Does not represent a physical clone"								
	<1..233								

```

/misc_feature
/note="P50017 ATP/GTP-binding site motif A (P-loop)"
complement(2873..2877)
/note="possible RBS upstream of RV2737c"
complement(3247..3453)

misc_feature
/note="P500321 reca signature"
complement(2644..2667)
/gene="reca"

misc_feature
/note="P500321 reca signature"
complement(2197..2223)
/gene="reca"

misc_feature
/note="reca intei1"
complement(790..807)
/gene="reca"

misc_feature
/note="P300881 Prefaji splicing signature"
complement(2197..2223)
/gene="reca"

misc_feature
/note="P500321 reca signature"
complement(2644..2667)
/gene="reca"

misc_feature
/note="P500017 ATP/GTP-binding site motif A (P-loop)"
complement(2873..2877)
/note="possible RBS upstream of RV2737c"
complement(3247..3453)

gene
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
/clone="Y154"
complement(3..527)
/gene="recX"
complement(3..527)
/note="RV2736c" (MTV002.01c), len: 174 aa. recX, similar
to eg. RECK_PSEAE_P37860 regulatory protein recX from
Pseudomonas aeruginosa (153 aa), fasta scores; opt: 161
z-score: 257.2 E(): 3.6e-07, 30.7% identity in 137 aa
overlap. Overlaps and extends CDS from overlapping consid
MTCY154.16c"
/codon_start=1
/transl_table=1
/product="recX"
/protein_id="CA15532.1"
/db_xref="GI:2624257"
/db_xref="SPTREMBL:Q33280"
/translation="MTVSCPPSPSTSESEDEOARALCLTLTARSRTBAIAGOLAKRKY
PEDIGNAGEERGAKEKLYRRARLRREVLIDDSIDEARVSRRLVAMLARIGQITLACEY
VIALDAERERRRV"
complement(493..2865)
/gene="reca"
complement(493..2865)
/gene="reca"
/note="RV2737c" (MTV002.02c), len: 790 aa. recA,
identical to RECA_MYCTU_P26345 recA protein (790 aa).
Contains self-splicing protein element (intein) from 2294
to 969 (c), similar to intein II from TR:E332317
(EMBL:Y13030) DNA-directed DNA polymerase (EC 2.7.7.7) from
Thermococcus sp. (1829 aa), fasta scores; opt: 81 z-score:
235.2 E(): 5e-05, 24.6% identity in 183 aa overlap.
contains P500017 ATP/GTP-binding site motif A (P-loop),
P500321 recA signature, and P500881 Protein splicing
signature. See Davis et al, (1992) Cell71(2):201-210"
/codon_start=1
/transl_table=1
/product="reca"
/protein_id="CA15533.1"
/db_xref="GI:2624258"
/db_xref="SWISS-PROT:P26345"
/translation="MTQTPDREKALELVAQIEKSYGKGSYMRIGDEARQPISYIPTG
SIALDVAIGICGLPGRAVIEIYGPSSGKTTVALHANAQAAGVAAFIDAEHALPD
DYAKKIDGALVLSOPDDEGOALEINDMIRSGALDIYVDSVALVPAELEGKE
GDSYVGOAKRMSOALKRMTGALNNSGCTAFINOLBDKIGMGSGPETHTGKALKE
YASVMDRREVEITLKDGTNAVGNRTKRVKVAAGKCLAESTRIFDPYVTGTHTEIDVDS
RKLPHVAAADGTLHARPAVSWFDQGRDVIYGLRAGALVMAIPHKVLTETGERRA
AGELEKGRVQVAPRFDFGSDSPAPADHAILLYLIGDGDGVAGGKTPIINFVOR
ALIDVYTIATLCAAPDOGRISLAIAHRPGERNGVADLCOAGIYGLAKMTPIPI
WFEFDYDIADLVGNLFLFELFESDGVNSEOTGALRVGTTTSEOLAHIMHLRFGV
GSTRVDYPTOKRPSIVNGRLOSROKROVEFVRISGMNVTAFAESVMMGRGALITIO
ALPEPTQGRRGSSQATYILAAEMTDVAVLYLDERGYTQAEAAATGASGDPGRGKMOY
LGASLRDRQVADLADLDKDFLHDMLEELVSTIRVLTFRARTFDLVEBELTL
VAEGYVHNCSPPEKQAEFDILGKISREGSLIDMGVDQDLIRKSGAMFYEEBQQL
QGEKVARNFLEVNADVADAEIKRTKEKLGIAVVTDDPSNDGVLPAPYDF"
complement(790..2113)
/gene="reca"

misc_feature
/note="reca intei1"
complement(790..807)
/gene="reca"

misc_feature
/note="P300881 Prefaji splicing signature"
complement(2197..2223)
/gene="reca"

misc_feature
/note="P500321 reca signature"
complement(2644..2667)
/gene="reca"

misc_feature
/note="P500017 ATP/GTP-binding site motif A (P-loop)"
complement(2873..2877)
/note="possible RBS upstream of RV2737c"
complement(3247..3453)

```

```

CDS
/gene="Rv2738c"
/complement(3247..3453)
/note="Rv2738c"
N-terminus is highly similar to the N-terminus of the
upstream ORF RV002.07c (78.4% identity in 37 aa overlap);
also similar to AL020958|SC4H8.5 Streptomyces coelicolor
cosmid 4H8 (64aa) opt: 185 z-score: 283.5 E(): 2.9e-08;
39.7% identity in 63 aa overlap"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv2738c"
/protein_id="CA15534.1"
/db_xref="GI:2624260"
/db_xref="SPTREMBL:O33281"
/transl_table="MAGVRLTEFHERALHFGAAYGSSVLDHVLTFGDSRAQAQAI
EDGVEPRVWALCADFVPHDRW"
/complement(3464..4630)
/gene="Rv2739c"
/complement(3464..4630)
/note="Rv2739c" (MTV002.04c) len: 388 aa. Probable
transferase, similar to eg. TR-051560 (EMBL:L28170)
Pseudomonas aeruginosa rhamnosyl transferase (426 aa)
fasta scores; opt: 178 z-score: 226.3 E(): 1.9e-05; 25.9%
identity in 425 aa overlap. Equivalent to Mycobacterium
leprae protein MLCB33.02c (392 aa); fasta scores
gpl294723|MLCB33-2 Mycobacterium leprae cosmid B33 opt:
211.2 z-score: 2364.5 E(): 0; 80.9% identity in 388 aa
overlap"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv2739c"
/protein_id="CA15535.1"
/db_xref="GI:2624261"
/db_xref="SPTREMBL:O33282"
/transl_table="MRAVAVAGPPGHSFPAICQFRRAADPTLTGVELEAAR
AAGIDAVELDLATRDLDGARIHRRRAAVALNVRALPELIVSDVITACGG
MAELIDIPWELNPHLPIDGSKLPIDGSLPGLI RGLRDTATMALGRMWRG
LRORAAVVEIGLPARDPGLRLIATLPALFVRRPMPAAVVGPLHFPDTRVA
IPAGTGVVAVASTALDTGALTEVALQSITPGETVSGSRVLSGADLTPVPA
AVAGISQAEILTEADLVICGGHGVAKTLLAIPVAVVVGSGDQEIARVROGS
AVLRPLTADLVAAVNEVSSPFRFRAARAAASVAGADPVRVCHDALALAC"
/4674..5123
/gene="Rv2740"
/4674..5123
/note="Rv2740" (MTV002.05) len: 149 aa. unknown, similar
to M. leprae hypothetical protein TR-049650 (EMBL:294723)
MLCB33.03 (B2235_F3_140) (178 aa) fasta scores; opt: 498
z-score: 538.9 E(): 7.3e-23; 51.6% identity in 161 aa
overlap"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv2740"
/protein_id="CA15536.1"
/db_xref="GI:2624262"
/db_xref="SPTREMBL:O33283"
/transl_table="MAELTSPETP TTEAIRAVLAFNALQNEDEFDVDAAGDL
VYENVSRIRGRRATILRRMGKRGVFKIRIDAGAAVLTERTDALIIGPLAV
OFWYCVFENDGRITIMRDYFDYDMFKGLRLVALVPSLAKATL"
/5355..6932
/gene="PE_PGRS"
/5355..6932
/note="PE_PGRS"
tuberculosis PE_PGRS subfamily, similar to many eg.

Query Match 99.9% Score 1234.4; LB 1; Length 56414;
Best Local Similarity 99.9% Pred. No. 6.9e-170;
Matches 1235; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATCTTGCAGATTATTCGAACCTTCTTCACACGAGCGTACGATTCGAGAGGGGTAATC 60

```

Db 39281 AAGGCTTTGACGACGACGAGGGGCGTACTGTCGACGAGGGTGGCCACCGACTGGG 39340
 QY 1141 GTGCGCTTACACGACCGCCGACAGTGTCTGCTGACTCTCGGCGGTGCTTACGCGGAG 1200
 Db 39341 GTGCGCTTACACGACCGCCGACAGTGTCTGCTGACTCTCGGCGGTGCTTACGCGGAG 39400
 QY 1201 CACACGTGGAGTAAGGAGGAGTGTGCGGC 1235
 Db 39401 CACACGTGGAGTAAGGAGGAGTGTGCGGC 39436

RESULT 6
 A87606 1235 bp DNA PAT 22-JAN-2000
 LOCUS Sequence 3 from Patent WO9836089.
 DEFINITION A87606
 ACCESSION A87606.1 GI:6736246
 VERSION
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.

REFERENCE 1 (bases 1 to 1235)
 Flohe, L. and Singh, M.
 TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
 TITLE Patent: WO 9836089-A 20-AUG-1998;
 JOURNAL FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source 1..1235
 Location/Qualifiers
 /organism="unidentified"
 /db_xref="taxon:32644"

BASE COUNT 236 a 394 c 382 g 220 t 3 others
 ORIGIN

Query Match 99.7% Score 1232: DB 5: Length 1235;
 Best Local Similarity 99.8%; Pred. No. 3.5e-159;
 Matches 1232: Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATCTGACAGATTAAATGCAACTTCTCTACACTGAGCGCTAGCATGATGAGAGGGTAAATC 60
 Db 1 ATCTGACAGATTAAATGCAACTTCTCTACACTGAGCGCTAGCATGATGAGAGGGTAAATC 60
 QY 61 ATGCGCGTGTATTCGACGACGAGACCAAAACAAGAAATTCGGGTGGCATACCCCG 120
 Db 61 ATGCGCGTGTATTCGACGACGAGACCAAAACAAGAAATTCGGGTGGCATACCCCG 120
 QY 121 GCGGCGTGGGGAATACCGCGTGTGCGCATGAGTGTCTCATCGAGGAGTCCGGA 180
 Db 121 GCGGCGTGGGGAATACCGCGTGTGCGCATGAGTGTCTCATCGAGGAGTCCGGA 180
 QY 181 GAGGCGTGGCTATCACGACGCGGATTTCAAGCGGAGCGGCAACTGGTGGCAC 240
 Db 181 GAGGCGTGGCTATCACGACGCGGATTTCAAGCGGAGCGGCAACTGGTGGCAC 240
 QY 241 GCGGACAGTGTGGGCGGAGCGCTATTTATGTCAGTGTAAAGAACCATAGCGGG 300
 Db 241 GCGGACAGTGTGGGCGGAGCGCTATTTATGTCAGTGTAAAGAACCATAGCGGG 300
 QY 301 GAATACGCGCGCTGCGACAGCGGAGATCTTGTTCAGTTCTTGCATTGGCGCGTCA 360
 Db 301 GAATACGCGCGCTGCGACAGCGGAGATCTTGTTCAGTTCTTGCATTGGCGCGTCA 360
 QY 361 CGTGTTCACAGATGCGTGTGTGATTCGCGACACGCTCAATGTGCTACAGACGCG 420
 Db 361 CGTGTTCACAGATGCGTGTGTGATTCGCGACACGCTCAATGTGCTACAGACGCG 420
 QY 421 CAGACCGCGAGCGGCGACATACCTGCTTGCCTGATGAGGAGGAGTGGCGGTGACTC 480
 Db 421 CAGACCGCGAGCGGCGACATACCTGCTTGCCTGATGAGGAGGAGTGGCGGTGACTC 480
 QY 481 GCGGCGCGAGTGGGCGTTCACAGCTGATGCAACCAAGGCGGCGGCGGTGCGTGAATG 540
 Db 481 GCGGCGCGAGTGGGCGTTCACAGCTGATGCAACCAAGGCGGCGGCGGTGCGTGAATG 540

QY 541 GCGGCGGTGCGGCGCTGCAACCGCGGACGCTGTGTGATGCGGCGCGGACCGCGGC 600
 Db 541 GCGGCGGTGCGGCGCTGCAACCGCGGACGCTGTGTGATGCGGCGCGGACCGCGGC 600
 QY 601 TACACGACGCGCGCATGCGCAACGCGATGCGGCGGACCGCTTACGATTACATCAAC 660
 Db 601 TACACGACGCGCGCATGCGCAACGCGATGCGGCGGACCGCTTACGATTACATCAAC 660
 QY 661 ATGCAAACTTGGGACATCGACGCGGAGTTCTGGGCGGATCCACACTGCGCACTCA 720
 Db 661 ATGCAAACTTGGGACATCGACGCGGAGTTCTGGGCGGATCCACACTGCGCACTCA 720
 QY 721 TCGGCTACAGCTTCAGAGGTGCGCTCAACGTCGCCACTGCTGATTTGGGCGCTCTG 780
 Db 721 TCGGCTACAGCTTCAGAGGTGCGCTCAACGTCGCCACTGCTGATTTGGGCGCTCTG 780
 QY 781 GTGCGAGCGCGCAAGGACCCCAATTAATGCTGATTCGATTTGCTGCGGCTATGAACCA 840
 Db 781 GTGCGAGCGCGCAAGGACCCCAATTAATGCTGATTCGATTTGCTGCGGCTATGAACCA 840
 QY 841 GTGCGCTACGCTTCAGAGGTGCGCTCAACGTCGCCACTGCTGATTTGGGCGCTCTG 900
 Db 841 GTGCGCTACGCTTCAGAGGTGCGCTCAACGTCGCCACTGCTGATTTGGGCGCTCTG 900
 QY 901 ACCACCTACGACCAACCGGACGCTTCCGCTGACGACAGCTGTTTACTGCTGCGGAAC 960
 Db 901 ACCACCTACGACCAACCGGACGCTTCCGCTGACGACAGCTGTTTACTGCTGCGGAAC 960
 QY 961 ATGCGCGCTGCTGGTGGCGAAGAGTGTGACCTAGCGCTGACCAAGCGGATGCGCTAT 1020
 Db 961 ATGCGCGCTGCTGGTGGCGAAGAGTGTGACCTAGCGCTGACCAAGCGGATGCGCTAT 1020
 QY 1021 GTGCTGAGCTTGGCGACATGCTGCGGCGCGCTGCGCTGCAATCGGACCTAGGC 1080
 Db 1021 GTGCTGAGCTTGGCGACATGCTGCGGCGCGCTGCGCTGCAATCGGACCTAGGC 1080
 QY 1081 AAGGCTTTTCGACGACGCAAGGCGCTTACTGTCCGACAGGTTGGCCACCGACTGGGG 1140
 Db 1081 AAGGCTTTTCGACGACGCAAGGCGCTTACTGTCCGACAGGTTGGCCACCGACTGGGG 1140
 QY 1141 GTGCGCTTACACGACCGCGGAGGCTGCTGCTGACTCTGCGCGCTGCTTACGCGGAG 1200
 Db 1141 GTGCGCTTACACGACCGCGGAGGCTGCTGCTGACTCTGCGCGCTGCTTACGCGGAG 1200
 QY 1201 CACACGTGGGAGTAAGCAACGATGATGTCGCGC 1235
 Db 1201 CACACGTGGGAGTAAGCAACGATGATGTCGCGC 1235

RESULT 7
 A89746 1235 bp DNA PAT 22-JAN-2000
 LOCUS Sequence 3 from Patent WO9832862.
 DEFINITION A89746
 ACCESSION A89746.1 GI:6738280
 VERSION
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.

REFERENCE 1 (bases 1 to 1235)
 Flohe, L. and Singh, M.
 L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
 TITLE Patent: WO 9832862-A 30-JUL-1998;
 JOURNAL FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source 1..1235
 Location/Qualifiers
 /organism="unidentified"
 /db_xref="taxon:32644"

BASE COUNT 236 a 394 c 382 g 220 t 3 others
 ORIGIN

Query Match 99.7%; Score 1222; DB 5; Length 1235;
 Best Local Similarity 99.8%; Pred. No. 3.5e-169;
 Matches 1227; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 ATCTTGAGATTAATGAACTTTCTTCTACACTGAAAGGTAACAGATATGAGAGGGTAATC 60
DB 1 ATCTTGAGATTAATGAACTTTCTTCTACACTGAAAGGTAACAGATATGAGAGGGTAATC 60
QY 61 ATGGGCTGCGTATTCGAGCAGAGACCAAAACAAAGATTCGGGGTGGCATACCCCG 120
DB 61 ATGGGCTGCGTATTCGAGCAGAGACCAAAACAAAGATTCGGGGTGGCATACCCCG 120
QY 121 GCCGGCTGCGGAACTAACCCGTGTGGCATGAGGTGCTCATCAGGAGGTGCCGA 180
DB 121 GCCGGCTGCGGAACTAACCCGTGTGGCATGAGGTGCTCATCAGGAGGTGCCGA 180
QY 181 GAGGGCTGCGGTAATCAACGAGCGGATTTCAAGCGGCGGCAAGCGATGTCGGACC 240
DB 181 GAGGGCTGCGGTAATCAACGAGCGGATTTCAAGCGGCGGCAAGCGATGTCGGACC 240
QY 241 GCCGACCAAGTGTGGCGGCGGAGTATTTATGCTCAAGGTCAAAAGAACGATAGCGCG 300
DB 241 GCCGACCAAGTGTGGCGGCGGAGTATTTATGCTCAAGGTCAAAAGAACGATAGCGCG 300
QY 301 GAATACGCGCGCTGCGACACAGGCGAGATCTGTTCACGTTTGTGCAATTTGGCCGCTCA 360
DB 301 GAATACGCGCGCTGCGACACAGGCGAGATCTGTTCACGTTTGTGCAATTTGGCCGCTCA 360
QY 361 CGATCTTGACAGATGCTGTGTGATTCGGGCGACAGTCAATTTGCTTACAGAGACGCTC 420
DB 361 CGATCTTGACAGATGCTGTGTGATTCGGGCGACAGTCAATTTGCTTACAGAGACGCTC 420
QY 421 CAGACCGCGGAGCGGAGTACCCCTGTTGCCCGATGAGCGAAGTGGCCGCTGACTC 480
DB 421 CAGACCGCGGAGCGGAGTACCCCTGTTGCCCGATGAGCGAAGTGGCCGCTGACTC 480
QY 481 GCCGCGCGAGTGTGGCTTACACCTGATGAGAACCCAAAGGGGCGCGGTGTCTATG 540
DB 481 GCCGCGCGAGTGTGGCTTACACCTGATGAGAACCCAAAGGGGCGCGGTGTCTATG 540
QY 541 GCGGCGGTGCGCGGCTGCAACCGGCGGAGTGTGGTGTGATGCGGCGGCGGCGGCG 600
DB 541 GCGGCGGTGCGCGGCTGCAACCGGCGGAGTGTGGTGTGATGCGGCGGCGGCGGCGGCG 600
QY 601 TACAACGAGCGGCGATCGCCAAAGGAGTGGCGGAGCGGTTAGCGTTTACAGATCAAC 660
DB 601 TACAACGAGCGGCGATCGCCAAAGGAGTGGCGGAGCGGTTAGCGTTTACAGATCAAC 660
QY 661 ATGCAACAACTTGGGCAACTGAGCGGAGTGTGGCGGAGTGTGACACTGCTACTCA 720
DB 661 ATGCAACAACTTGGGCAACTGAGCGGAGTGTGGCGGAGTGTGACACTGCTACTCA 720
QY 721 TCGGCTACGAGTGTGAGGAGTGTGCAAAAGTGTGCAAGTGTGATTTGGGCGGCTGTG 780
DB 721 TCGGCTACGAGTGTGAGGAGTGTGCAAAAGTGTGCAAGTGTGATTTGGGCGGCTGTG 780
QY 781 GTGCGAGCGGCGCAAGGCAACCAAAATAGTGTGCAATTTGCTGCGGCAATTAACA 840
DB 781 GTGCGAGCGGCGCAAGGCAACCAAAATAGTGTGCAATTTGCTGCGGCAATTAACA 840
QY 841 GGTGCGGTGTGAGTGTGATGAGTGTGAGGAGGCGGCTGTTCGAAAGGTTCACGACCG 900
DB 841 GGTGCGGTGTGAGTGTGATGAGTGTGAGGAGGCGGCTGTTCGAAAGGTTCACGACCG 900
QY 901 ACCACCTACGACACCGGAGTGTGCGGTGAGACAGAGTGTGTTTACGTTGAGGCGGAC 960
DB 901 ACCACCTACGACACCGGAGTGTGCGGTGAGACAGAGTGTGTTTACGTTGAGGCGGAC 960
QY 961 ATGCGCGCTGTGAGGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1020
DB 961 ATGCGCGCTGTGAGGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1020
QY 1021 GTGCTGAGCTTGGCGGACATGCTGTGGGCGGCGGCTGTGGATTCGGGCACTAGCC 1080

```

```

DB 1021 GTGCTGAGCTTGGCGGACATGCTGTGGGCGGCGGCTGTGGATTCGGGCACTAGCC 1080
QY 1081 AAGGCTTTTCGACGACGACGAGGCGGTACTGTCCGAACGGGTGGCCGACGCTGGGG 1140
DB 1081 AAGGCTTTTCGACGACGACGAGGCGGTACTGTCCGAACGGGTGGCCGACGCTGGGG 1140
QY 1141 GTGCGCTTACGAGCGGCGGCGGAGTGTGCTGACTGTGCGGCTGTGCTGCTGCTGCTG 1200
DB 1141 GTGCGCTTACGAGCGGCGGCGGAGTGTGCTGACTGTGCGGCTGTGCTGCTGCTGCTG 1200
QY 1201 CACAGCTGCGGAGTAAAGGAGGAGTGTGCTG 1235
DB 1201 CACAGCTGCGGAGTAAAGGAGGAGTGTGCTG 1235
RESULT 8
A87608 1228 bp DNA PAT 22-JAN-2000
LOCUS A87608 Sequence 5 from Patent WO9836089.
DEFINITION A87608
ACCESSION A87608.1 GI:5736248
VERSION A87608.1 GI:5736248
KEYWORDS
SOURCE
ORGANISM
unidentified.
unclassified.
REFERENCE
1 (bases 1 to 1228)
AUTHORS
Flohe, L. and Singh, M.
TITLE
TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL
Patent: WO 9836089-A 20-AUG-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
source
1..1228
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 236 a 391 c 382 g 219 t
ORIGIN
Query Match 99.2%; Score 1226.4; DB 5; Length 1228;
Best Local Similarity 99.8%; Pred. No. 2.3e-168;
Matches 1227; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATCTTGAGATTAATGAACTTTCTTCTACACTGAAAGGTAACAGATATGAGAGGGTAATC 60
DB 1 ATCTTGAGATTAATGAACTTTCTTCTACACTGAAAGGTAACAGATATGAGAGGGTAATC 60
QY 61 ATGCGCGTGTGATTTGCAATTAATGAAAGTGTGGTGTGCTGCTGCTGCTGCTGCTGCTG 120
DB 61 ATGCGCGTGTGATTTGCAATTAATGAAAGTGTGGTGTGCTGCTGCTGCTGCTGCTGCTG 120
QY 121 GCCGGCTGCGGAGTGTGAGTGTGCAAAAGTGTGCAAGTGTGATTTGGGCGGCTGTG 180
DB 121 GCCGGCTGCGGAGTGTGAGTGTGCAAAAGTGTGCAAGTGTGATTTGGGCGGCTGTG 180
QY 181 GAGGGCTGCGGCTTACACGAGCGGAGTGTCAAGGCGGAGCGGCGGCACTGTGCGCAC 240
DB 181 GAGGGCTGCGGCTTACACGAGCGGAGTGTCAAGGCGGAGCGGCGGCACTGTGCGCAC 240
QY 241 GCGGAGCGTGTGAGGCGGAGTGTGATTTATGCTGTAAGTCAAGTCAAGTCAAGTCAAG 300
DB 241 GCGGAGCGTGTGAGGCGGAGTGTGATTTATGCTGTAAGTCAAGTCAAGTCAAGTCAAG 300
QY 301 GAATACGCGCGCTGCGACACAGGCGGAGTGTGCAAGTGTGCAAGTGTGATTTGGGCGG 360
DB 301 GAATACGCGCGCTGCGACACAGGCGGAGTGTGCAAGTGTGCAAGTGTGATTTGGGCGG 360
QY 361 GTGCTTTCGACGAGTGTGTTGATTCGGGCGGAGTGTGATTCGGGCGGAGTGTGATTCGG 420
DB 361 GTGCTTTCGACGAGTGTGTTGATTCGGGCGGAGTGTGATTCGGGCGGAGTGTGATTCGG 420
QY 421 CAGACCGCGGAGGCGGAGTGTGATTCGGGCGGAGTGTGATTCGGGCGGAGTGTGATTCGG 480

```


Db	421	CAGACGCGCAGCGGCGACTACCCCTGTTGGCCCGATGAGCGAAGTGC	CGCGTCACTC	480
Qy	481	GGCGCCAGTTGGCGCTTACCACTGATGCGAACCCAAAGGGGGCGGGTGTCTATG	540	
Db	481	GGCCCGCCAGGTTGGCGCTTACCACTGATGCGAACCCAAAGGGGGCGGGTGTCTATG	540	
Qy	541	GGGGGGGTTGGCCGGCGTGCMAACGGCGCGAAGTGTGTATGGGGCGCGACCGCGGC	600	
Db	541	GGCGGGGTGCCCGCGCTGCMAACGGCGCGAAGTGTGTATGGGGCGCGACCGCGGC	600	
Qy	601	TACAACGACGCGCGCATGCGCAACGGCGATGGCGCGACCGTTACGGTTCTAGACATCAAC	660	
Db	601	TACAACGACGCGCGCATGCGCAACGGCGATGGCGCGACCGTTACGGTTCTAGACATCAAC	660	
Qy	661	ATCGACAACTTGGCGCAACTTCGACAGCGCGAGTTGTGGGGCGGATCCAACTCGCTACTCA	720	
Db	661	ATCGACAAACTTGGCGCAACTTCGACAGCGCGAGTTGTGGGGCGGATCCAACTCGCTACTCA	720	
Qy	721	TCGGCTACGAGCTCGAGGGTGGCGCTCAAACTGCGCGACTGGTGTATGGGGCGCTCTG	780	
Db	721	TCGGCTACGAGCTCGAGGGTGGCGCTCAAACTGCGCGACTGGTGTATGGGGCGCTCTG	780	
Qy	781	GTGCGAGGCGCCAAAGGACCCAAATTATGTCGAAATTCCTTGTGGCGCATATGAACA	840	
Db	781	GTGCGAGGCGCCAAAGGACCCAAATTATGTCGAAATTCCTTGTGGCGCATATGAACA	840	
Qy	841	GGTGGGATAGTGAGATATAGCCATGACAGAGGGGGGCTTGCAAGGCTCAGACACG	900	
Db	841	GGTGGGATAGTGAGATATAGCCATGACAGAGGGGGGCTTGCAAGGCTCAGACACG	900	
Qy	901	ACCACTACGACCAACCCGAGCTTGCGCGTGCAGACAGCGCTGTTTATCTGCTGGCGAAC	960	
Db	901	ACCACTACGACCAACCCGAGCTTGCGCGTGCAGACAGCGCTGTTTATCTGCTGGCGAAC	960	
Qy	961	ATGCCCCCTTGGTGGCCGAAGAGCTGCACCTACGCGCTGACCAAGCGAGATGCCGAT	1020	
Db	961	ATGCCCCCTTGGTGGCCGAAGAGCTGCACCTACGCGCTGACCAAGCGAGATGCCGAT	1020	
Qy	1021	GTGCTCGAGTTCGCGACCAATGAGCTGGCGGGCGCGTGGCGTGCATTCGGCACTAGCC	1080	
Db	1021	GTGCTCGAGTTCGCGACCAATGAGCTGGCGGGCGCGTGGCGTGCATTCGGCACTAGCC	1080	
Qy	1081	AAAGGCTTTTCGAGCGACGAAGAGGGGGCTTACTGTCCGAACGGGTGGCCACACCTGGGG	1140	
Db	1081	AAAGGCTTTTCGAGCGACGAAGAGGGGGCTTACTGTCCGAACGGGTGGCCACACCTGGGG	1140	
Qy	1141	GTGGCGTTCACGAGCCCGCGAGGCTGTGACCTGTGGCGCTGTGGCGCTGTACGCCGAG	1200	
Db	1141	GTGGCGTTCACGAGCCCGCGAGGCTGTGACCTGTGGCGCTGTGGCGCTGTACGCCGAG	1200	
Qy	1201	CACACGTGCGGATTAAGGGAAGCGATGA	1228	
Db	1201	CACACGTGCGGATTAAGGGAAGCGATGA	1228	
RESULT	9			
LOCUS	A89748	1228 bp	DNA	
DEFINITION	Sequence 5 from Patent WO9832862.		PAT	
ACCESSION	A89748		22-JAN-2000	
VERSION	A89748.1	GI:6738282		
KEYWORDS				
SOURCE	unidentified.			
ORGANISM	unclassified.			
REFERENCE	1 (bases 1 to 1228)			
AUTHORS	Flohe, L. and Singh, M.			
TITLE	L-LANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM			
JOURNAL	Patent: WO 9832862-A-30-JUL-1998;			
FEATURES	FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)			
source	Location/Qualifiers			
	1..1228			
	/organism="unidentified"			


```

QY 961 ATGCCGCTCGGTCGGAAGAGCTGACCTACGCGCTACCAACGCGAGCATGCGCTAT 1020
DB 961 ATGCCGCTCGGTCGGAAGAGCTGACCTACGCGCTACCAACGCGAGCATGCGCTAT 1020
QY 1021 GTGCTCGACCTGCGGACCATGCTGCGGCGGCGTGCCTGCAATCCGCGACTAGCC 1080
DB 1021 GTGCTCGACCTGCGGACCATGCTGCGGCGGCGTGCCTGCAATCCGCGACTAGCC 1080
QY 1081 AAGGCTTTCGACGCGAGGAGGCGGCTGCTGCTGCAACCTGCGGCGGCGGCGG 1140
DB 1081 AAGGCTTTCGACGCGAGGAGGCGGCTGCTGCTGCAACCTGCGGCGGCGGCGG 1140
QY 1141 GTGCGCTTCACGAGCGCGCCACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
DB 1141 GTGCGCTTCACGAGCGCGCCACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
QY 1201 CACACGCTCGGAGTAAGGAGGAGGATGA 1228
DB 1201 CACACGCTCGGAGTAAGGAGGAGGATGA 1228

```

```

RESULT 10
A87609
LOCUS A87609 1235 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 6 from Patent WO9836089.
ACCESSION A87609
VERSION A87609.1 GI:6736249
KEYWORDS
SOURCE .
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1235)
AUTHORS Flohe, L. and Singh, M.
TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE
JOURNAL Patent: WO 9836089-A-20-AUG-1998;
FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
FEATURES
source 1. 1235
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 235 a 395 c 384 g 220 t 1 others
ORIGIN

```

```

Query Match 98.9%; Score 1223; DB 5; Length 1235;
Best Local Similarity 99.8%; Pred. No. 7e-168;
Matches 1234; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

```

QY 1 ATCTTCAGATTATATGAACTTTCTTACACGTAAGCGTACAGTATGAGAGGGTAAATC 60
DB 1 ATCTTCAGATTATATGAACTTTCTTACACGTAAGCGTACAGTATGAGAGGGTAAATC 60
QY 61 ATGCGGCTGGGATTCGAGCGGAGCAAAACAAAGAAATTCGGGCTGCGCATACCCCG 120
DB 61 ATGCGGCTGGGATTCGAGCGGAGCAAAACAAAGAAATTCGGGCTGCGCATACCCCG 120
QY 121 GCGGCGTGGCGGAATTAACCGCTGCTGCGCATGAGTCTCATCCAGGAGGCGGGA 180
DB 121 GCGGCGTGGCGGAATTAACCGCTGCTGCGCATGAGTCTCATCCAGGAGGCGGGA 180
QY 181 GAGGCTCGGCTATACCAACGCGGATTCGAAGGCGGC,GGCGCGCACTGTGCGGAC, 240
DB 181 GAGGCTCGGCTATACCAACGCGGATTCGAAGGCGGC,GGCGCGCACTGTGCGGAC, 240
QY 241 GCGGACAGAGTGGGCGGCGGCGGCTGATTTATGCTCAAGGTCAAAACGATAGCGGCG 300
DB 241 GCGGACAGAGTGGGCGGCGGCGGCTGATTTATGCTCAAGGTCAAAACGATAGCGGCG 300
QY 301 GATAAGCGCGCTGCGAGACGCGGAGATCTTTCATCTTTCGATTTGCGCGGCTCA 360
DB 301 GATAAGCGCGCTGCGAGACGCGGCGGAGATCTTTCATCTTTCGATTTGCGCGGCTCA 360
QY 361 CGTGCTTGACCGAGTGGTGTGGATTCGCGGACCAAGCTCAATTGCTTACGAGACCGTC 420

```

```

DB 360 CGTGCTTGACCGAGTGGTGTGGATTCGCGGACCAAGCTCAATTGCTTACGAGACCGTC 419
QY 421 CAGACCGCGAGCGGCGCACTACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB 421 CAGACCGCGAGCGGCGCACTACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 481 GCGGCGGCTGGGCTTACCACTGATGGAACCCCAAGGCGGCGGCGGCTGCTGATG 540
DB 481 GCGGCGGCTGGGCTTACCACTGATGGAACCCCAAGGCGGCGGCGGCTGCTGATG 540
QY 541 GCGGCGGCTGGGCTGGAACCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 541 GCGGCGGCTGGGCTGGAACCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 601 TACAACGAGCGCGGCTATGCCAAGCGGCTGCGGCGGCGGCTTACGCTTACATCAAC 660
DB 601 TACAACGAGCGCGGCTATGCCAAGCGGCTGCGGCGGCGGCTTACGCTTACATCAAC 660
QY 661 ATGACAAACTTCGCACTGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB 661 ATGACAAACTTCGCACTGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 721 TCGGCTTACGAGCTGAGGCTGCGGCTCAACGCTGCGGCTGCTGCTGCTGCTGCTGCTG 780
DB 721 TCGGCTTACGAGCTGAGGCTGCGGCTCAACGCTGCGGCTGCTGCTGCTGCTGCTGCTG 780
QY 781 GTGCGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
DB 781 GTGCGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
QY 841 GGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
DB 841 GGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 901 ACCACTTACGACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
DB 901 ACCACTTACGACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
QY 961 ATGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
DB 961 ATGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 1021 GTGCTGAGCTTCCGCGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
DB 1021 GTGCTGAGCTTCCGCGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
QY 1081 AAGGCTTTTCGACGCAAAACAAAGGCGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
DB 1081 AAGGCTTTTCGACGCAAAACAAAGGCGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
QY 1141 GTGCGCTTACCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
DB 1141 GTGCGCTTACCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
QY 1201 CACACGCTCGGAGTAAGGAGGAGGATGATGTCGCGC 1236
DB 1201 CACACGCTCGGAGTAAGGAGGAGGATGATGTCGCGC 1236

```

```

RESULT 11
A89749
LOCUS A89749 1235 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 6 from Patent WO9832862.
ACCESSION A89749
VERSION A89749.1 GI:6738283
KEYWORDS
SOURCE .
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1235)
AUTHORS Flohe, L. and Singh, M.
TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM

```

JOURNAL Patent: WO 9832862-A 30-JUL-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES Location/Qualifiers
 source 1.1235
 /db_xref="taxon:32644"
 BASE COUNT 235 a 395 c 384 g 220 t 1 others
 ORIGIN

Query Match 98.9%; Score 1223; DB 5; Length 1235;
 Best Local Similarity 99.8%; Pred. No. 7e-168;
 Matches 1234; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

OY 1 ATCTTGACATTAATGACACTTCTTCCACATGAAAGCTAGACATGAGAGGGTAATC 60
DB 1 ATCTTGACATTAATGACACTTCTTCCACATGAAAGCTAGACATGAGAGGGTAATC 60
OY 61 ATGCGGTCGTGATTCGACCGAGACCAAAACAAAGAAATTCGGGTGGCCATCACCCG 120
DB 61 ATGCGGTCGTGATTCGACCGAGACCAAAACAAAGAAATTCGGGTGGCCATCACCCG 120
OY 121 GCGGCGTCGCGGAAGTAACCCGTCGTGCGCATGAGTGCTATCCAGGAGGTGCCGA 180
DB 121 GCGGCGTCGCGGAAGTAACCCGTCGTGCGCATGAGTGCTATCCAGGAGGTGCCGA 180
OY 181 GAGGGTCGTGATACGACGCGGATTCGAAGGCGGAGCGCGGACATGGTGGGACCC 240
DB 181 GAGGGTCGTGATACGACGCGGATTCGAAGGCGGAGCGCGGACATGGTGGGACCC 240
OY 241 GCGGACAGAGTGTGGGCGGACGCTGATTTATGCTCAAGGTCAAAAGACGATAGCGGG 300
DB 241 GCGGACAGAGTGTGGGCGGACGCTGATTTATGCTCAAGGTCAAAAGACGATAGCGGG 300
OY 301 GAATACGCGCCCTGCGGACAGGCGAGATCTGTACGTTCTGTGATTTGGCCGCTCA 360
DB 301 GAATACGCGCCCTGCGGACAGGCGAGATCTGTACGTTCTGTGATTTGGCCGCTCA 360
OY 361 GCGGTCGACGAGTGGTGTGATTCGCGGACGAGCATTCCTACGAGACGCTC 420
DB 361 GCGGTCGACGAGTGGTGTGATTCGCGGACGAGCATTCCTACGAGACGCTC 420
OY 421 GCGGCGCGGACGCGGACATACCCCTGCTGCGGCGATGAGCGAAATGCGCGGTGACTC 480
DB 421 GCGGCGCGGACGCGGACATACCCCTGCTGCGGCGATGAGCGAAATGCGCGGTGACTC 480
OY 481 GCGGCGGAGTGGGCGCTTACACCTGATGCGAACCAAGGGGCGCGGTGCTGATG 540
DB 481 GCGGCGGAGTGGGCGCTTACACCTGATGCGAACCAAGGGGCGCGGTGCTGATG 540
OY 541 GCGGCGGAGTGGGCGCTTACACCTGATGCGAACCAAGGGGCGCGGTGCTGATG 599
DB 541 GCGGCGGAGTGGGCGCTTACACCTGATGCGAACCAAGGGGCGCGGTGCTGATG 599
OY 601 TACAAGCGACCGCGATCGCAACGCGATGGGCGGACCGGTTACGGTTTACGATCAAC 660
DB 601 TACAAGCGACCGCGATCGCAACGCGATGGGCGGACCGGTTACGGTTTACGATCAAC 660
OY 661 ATGCGCAAACTTGGGTAATCGACGCGGATTCGGGCGGATCCACACTGGCTACTA 720
DB 661 ATGCGCAAACTTGGGTAATCGACGCGGATTCGGGCGGATCCACACTGGCTACTA 720
OY 721 TCGGCGTACGAGCTGAGGGTGGCGTCAACGTCGCGACTGGTGTGAGGGGCGCTCTG 780
DB 721 TCGGCGTACGAGCTGAGGGTGGCGTCAACGTCGCGACTGGTGTGAGGGGCGCTCTG 780
OY 781 GTGCGAGGCGGCAAGGACCAAAATAGTCTGAAATTCATCTGTCCGCAATGAAACA 840
DB 781 GTGCGAGGCGGCAAGGACCAAAATAGTCTGAAATTCATCTGTCCGCAATGAAACA 840
OY 841 GTGCGAGTCTGATGATTAAGCATGACGAGGCGGCTGTTGCAAGGCTACGACCG 900
DB 841 GTGCGAGTCTGATGATTAAGCATGACGAGGCGGCTGTTGCAAGGCTACGACCG 900

```

```

OY 901 ACCACTACGACACCCGACGCTTCCGCGGACGACGACGCTGTTTACGCTGGCGAAC 960
DB 901 ACCACTACGACACCCGACGCTTCCGCGGACGACGACGCTGTTTACGCTGGCGAAC 960
OY 961 ATGCGCGCTCGGTGCGGAAAGAGTGCAGCTTACGCGCTGACCAAGCGAGTGGCTAT 1020
DB 961 ATGCGCGCTCGGTGCGGAAAGAGTGCAGCTTACGCGCTGACCAAGCGAGTGGCTAT 1020
OY 1021 GTGCTGACGCTTCCGACCATGCGTGGCGGCGGCGGCTGCGGATTCGATCCGACTACCC 1080
DB 1021 GTGCTGACGCTTCCGACCATGCGTGGCGGCGGCGGCTGCGGATTCGATCCGACTACCC 1080
OY 1081 AAGGCTCTTCAGCAGCAAGGAGGCTGCTGCTGCAAGGAGGAGGAGGAGGAGGAGG 1140
DB 1081 AAGGCTCTTCAGCAGCAAGGAGGCTGCTGCTGCAAGGAGGAGGAGGAGGAGGAGG 1140
OY 1141 GTGCGGTCACGAGCGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
DB 1141 GTGCGGTCACGAGCGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
OY 1201 CACAGCTCGGAGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1236
DB 1201 CACAGCTCGGAGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1236

```

RESULT 12

LOCUS A87611 1235 bp DNA PAT 22-JAN-2000
 DEFINITION Sequence 8 from Patent WO9836089.
 ACCESSION A8 611
 VERSION A8 311.1 GI:6736251
 KEYWORDS
 SOURCE
 ORGANISM
 unclassified.
 unclassified.
 unclassified.

REFERENCE 1 (bases 1 to 1235)

AUTHORS Flohe, L. and Singh, M.

TITLE TEST KIT FOR TUBERCULOSIS DIAGNOSIS OR THE LIKE

JOURNAL Patent: WO 9836089-A 20-AUG-1998;

FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)

FEATURES Location/Qualifiers

source 1.1235
 /db_xref="taxon:32644"
 /organism="unclassified"
 BASE COUNT 235 a 394 c 385 g 220 t
 ORIGIN

Query Match 98.9%; Score 1222.4; DB 5; Length 1235;
 Best Local Similarity 99.8%; Pred. No. 8.6e-168;
 Matches 1234; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

OY 1 ATCTTGACATTAATGACACTTCTTCCACATGAAAGCTAGACATGAGAGGGTAATC 60
DB 1 ATCTTGACATTAATGACACTTCTTCCACATGAAAGCTAGACATGAGAGGGTAATC 60
OY 61 ATGCGGTCGTGATTCGACCGAGACCAAAACAAAGAAATTCGGGTGGCCATCACCCG 120
DB 61 ATGCGGTCGTGATTCGACCGAGACCAAAACAAAGAAATTCGGGTGGCCATCACCCG 120
OY 121 GCGGCGTCGCGGAAGTAACCCGTCGTGCGCATGAGTGCTATCCAGGAGGTGCCGA 180
DB 121 GCGGCGTCGCGGAAGTAACCCGTCGTGCGCATGAGTGCTATCCAGGAGGTGCCGA 180
OY 181 GAGGGTCGTGATACGACGCGGATTCGAAGGCGGACGCGGCAACTGTGGGACCC 240
DB 181 GAGGGTCGTGATACGACGCGGATTCGAAGGCGGACGCGGCAACTGTGGGACCC 240
OY 241 GCGGACAGAGTGTGGGCGGACGCTGATTTATGCTCAAGGTCAAAAGACGATAGCGGG 300
DB 241 GCGGACAGAGTGTGGGCGGACGCTGATTTATGCTCAAGGTCAAAAGACGATAGCGGG 300

```

QY 301 GAATACGGCCGCTGCGACAGCGGAGATCTGTTACATCTTCTGATTTGGCCGCTCA 360
 DB 301 GATACGGCCGCTGCGACAGCGGC-GATCTTGTACGTTCTTGATTTGGCCGCTCA 359
 QY 361 CGTCTTGCACGATGCGCTTGTGATTCGGACACAGCTCAATTGCTACAGACCGTC 420
 DB 360 CGTCTTGCACGATGCGCTTGTGATTCGGACACAGCTCAATTGCTACAGACCGTC 419
 QY 421 CAGACCGCGACGCGGACATACCCCTGCTTGGCCGACATGAGGAATGCGCGCTGACTC 480
 DB 420 CAGACCGCGACGCGGACATACCCCTGCTTGGCCGACATGAGGAATGCGCGCTGACTC 479
 QY 481 GCGCCGACGATGCGGCTTACCACTGATGCAACCAAGGCGCGGCTGTGTATG 540
 DB 480 GCGCCGACGATGCGGCTTACCACTGATGCAACCAAGGCGCGGCTGTGTATG 539
 QY 541 GCGCGGCTGCGCGCTGCAACCGCGGACGCTGTGTATGCGCGCGGACCGCGG 600
 DB 540 GCGCGGCTGCGCGCTGCAACCGCGGACGCTGTGTATGCGCGCGGACCGCGG 599
 QY 601 TACAACGCGACCGGCTGCGACAGCGGATGGGCGGACGTTACGGTTGTACATCAAC 660
 DB 600 TACAACGCGACCGGCTGCGACAGCGGATGGGCGGACGTTACGGTTGTACATCAAC 659
 QY 661 ATGCAACAACTTGGCACTGCAAGCGGATTCGCGCGGATCCACACTGCTACTCA 720
 DB 660 ATGCAACAACTTGGCACTGCAAGCGGATTCGCGCGGATCCACACTGCTACTCA 719
 QY 721 TCGGCTTACGAGCTGAGGGTGGCTCAAAAGTGGCGGACCTGCTATG 780
 DB 720 TCGGCTTACGAGCTGAGGGTGGCTCAAAAGTGGCGGACCTGCTATG 779
 QY 781 GTGCGAGGCGCAAGGACCAAAATAGTCTGCAATTCCTGCTGCGATATGAACCA 840
 DB 780 GTGCGAGGCGCAAGGACCAAAATAGTCTGCAATTCCTGCTGCGATATGAACCA 839
 QY 841 GGTGCGTACTGTGTATAGCCATGACAGGCGGCTGTTTCGAAGGCTCAGACCG 900
 DB 840 GGTGCGTACTGTGTATAGCCATGACAGGCGGCTGTTTCGAAGGCTCAGACCG 899
 QY 901 ACCACTTACGACACCGGACGTTGCGGCTGACAGACGCTGTTTACGCTGGGGAAC 960
 DB 900 ACCACTTACGACACCGGACGTTGCGGCTGACAGACGCTGTTTACGCTGGGGAAC 959
 QY 961 ATGCGCGCTGCTGCGCAAGAGCTGACCTAGCGGCTGACCAACGCGAGATGCGTAT 1020
 DB 960 ATGCGCGCTGCTGCGCAAGAGCTGACCTAGCGGCTGACCAACGCGAGATGCGTAT 1019
 QY 1021 GTGCTGAGCTTGCAGCAATGCTGCGGCGGCTGCGGCTGCAATCGGCACTAGCC 1080
 DB 1020 GTGCTGAGCTTGCAGCAATGCTGCGGCGGCTGCGGCTGCAATCGGCACTAGCC 1079
 QY 1081 AAGGCTTTCGACGACGAAAGGCGGCTTACGTCGAACGCGTGGCCACGCGTGGGG 1140
 DB 1080 AAGGCTTTCGACGACGAAAGGCGGCTTACGTCGAACGCGTGGCCACGCGTGGGG 1139
 QY 1141 GTGCGTTCACGACGCGGCTGCTGCTGACTCTGCGGCTGCTGCTGCTGCTGCTG 1200
 DB 1140 GTGCGTTCACGACGCGGCTGCTGCTGACTCTGCGGCTGCTGCTGCTGCTGCTG 1199
 QY 1201 CACACGTGCGGAGTAAGGAAGGAGATGTCGGC 1236
 DB 1200 CACACGTGCGGAGTAAGGAAGGAGATGTCGGC 1235

RESULT 13
 LOCUS A89751 1235 bp DNA
 DEFINITION Sequence 8 from Patent WO9832862.
 ACCESSION A89751
 VERSION A89751.1 GI:6738285
 KEYWORDS
 SOURCE unidentified.

ORGANISM unidentified
 unclassified.
 REFERENCE 1 (bases 1 to 1235)
 AUTHORS Flohe, L. and Singh, M.
 TITLE L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
 JOURNAL Patent: WO 9832862-A 30-JUL-1998;
 FLOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 FEATURES
 source 1..1235
 location/Qualifiers
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 236 a 394 c 385 g 220 t
 ORIGIN
 Query Match 98.9%; Score 1222.4; DB 5; Length 1235;
 Best Local Similarity 99.8%; Pred. No. 8,6e-168;
 Matches 1234; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 1 ATCTTCAGATTAAATGCACTTCTTCACTGAAAGCTACAGTATGAGAGGGTAAATC 60
 DB 1 ATCTTCAGATTAAATGCACTTCTTCACTGAAAGCTACAGTATGAGAGGGTAAATC 60
 QY 61 ATGCGCTGCGTATTCGACCGAGACCAAAACAAGATTCGGGTCATACCCCG 120
 DB 61 ATGCGCTGCGTATTCGACCGAGACCAAAACAAGATTCGGGTCATACCCCG 120
 QY 121 GCGGCGTGGGGAACCTAACCCGTCGTGCGATGAGGTCATGCAAGGCGGAGCGGGA 180
 DB 121 GCGGCGTGGGGAACCTAACCCGTCGTGCGATGAGGTCATGCAAGGCGGAGCGGGA 180
 QY 181 GAGGCTGCGGTATCACCGACGCGGATTTCAAGCGGCGGCGCAACTGTCGGCAC 240
 DB 181 GAGGCTGCGGTATCACCGACGCGGATTTCAAGCGGCGGCGCAACTGTCGGCAC 240
 QY 241 GCGGACCAAGGTGTGGCGCGGACGCTGATTTATGCTCAAGTCAAGAACCGATAGCGCG 300
 DB 241 GCGGACCAAGGTGTGGCGCGGACGCTGATTTATGCTCAAGTCAAGAACCGATAGCGCG 300
 QY 301 GAATACGGCGCGCTGCGACAGCGGACAGATCTGTTCAAGTTCGATTTGGCGCGTCA 360
 DB 301 GAATACGGCGCGCTGCGACAGCGGACAGATCTGTTCAAGTTCGATTTGGCGCGTCA 359
 QY 361 CGTCTTGCACGATGCGCTTGTGATTCGGACACAGCTCAATTGCTACAGACCGTC 420
 DB 360 CGTCTTGCACGATGCGCTTGTGATTCGGACACAGCTCAATTGCTACAGACCGTC 419
 QY 421 CAGACCGCGACGCGGACACCTGCTTGCUCGATGAGGAAAGTGGCTGACTC 480
 DB 421 CAGACCGCGACGCGGACACCTGCTTGCUCGATGAGGAAAGTGGCTGACTC 479
 QY 481 GCGGCGCAGGTTGGGCTTACCACTGATGCGAACCAGAGGCGGCGGCTGTGTATG 540
 DB 480 GCGGCGCAGGTTGGGCTTACCACTGATGCGAACCAGAGGCGGCGGCTGTGTATG 539
 QY 541 GCGGCGGCTGCGCGCTGCAACCGCGGACGCTGCTGATGAGGCGGCGGACCGCGGC 600
 DB 540 GCGGCGGCTGCGCGCTGCAACCGCGGACGCTGCTGATGAGGCGGCGGACCGCGGC 599
 QY 601 TACAACGACCGCGGATGCGCAAGGCGATGGCGGACCGTTACGGTCTAGACTCAAC 660
 DB 600 TACAACGACCGCGGATGCGCAAGGCGATGGCGGACCGTTACGGTCTAGACTCAAC 659
 QY 661 ATGCAACAACTTGGCACTGCAAGCGGATTCGCGGCGGATCCACACTGCTACTCA 720
 DB 660 ATGCAACAACTTGGCACTGCAAGCGGATTCGCGGCGGATCCACACTGCTACTCA 719
 QY 721 TCGGCTTACGAGCTGAGGGTGGCTCAAAAGTGGCGACTGCTGATTTGGGCGCTGCTG 780
 DB 720 TCGGCTTACGAGCTGAGGGTGGCTCAAAAGTGGCGACTGCTGATTTGGGCGCTGCTG 779
 QY 781 GTGCGAGGCGGCAAGGACCAAAATAGTCTGCAATTCCTGCTGCGCATATGAACCA 840
 DB 781 GTGCGAGGCGGCAAGGACCAAAATAGTCTGCAATTCCTGCTGCGCATATGAACCA 840

DEFINITION Sequence 2 from Patent WO9832862.
 ACCESSION AB9745
 VERSION AB9745.1 GI:6738279
 KEYWORDS
 SOURCE
 ORGANISM
 unclassified.
 unclassified.
 unclassified.
 1 (bases 1 to 1245)
 Fiohe, L. and Singh, M.
 L-ALANINE DEHYDROGENASE OF MYCOBACTERIUM MARINUM
 TITLE
 JOURNAL
 FIOHE LEOPOLD (DE); SINGH MAHAVIR (DE)
 Location/Qualifiers
 1..1245
 /organism="unclassified"
 /db_xref="taxon:32644"
 BASE COUNT 238 a 398 c 387 g 222 t
 ORIGIN

Query Match 98.7%; Score 1220; DI 5; Length 1245;
 Best Local Similarity 99.5%; Pred. No. 1,9e-167;
 Matches 1236; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

QY 1 ATCTGCGAGTATATCGAATCTTCTTCACTGAGCGTACGATCGAGAGGGTAAATC 60
 Db 1 ATCTTGCAGATTAATCGAATCTTCTTCACTGAGCGTACGATCGAGAGGGTAAATC 60
 QY 61 ATGCGCGTGTATTCGACGAGACCAAAACG-----AATTCGGGTGGCCATC 114
 Db 61 ATGCGCGTGTATTCGACGAGACCAAAACGAAATTCGAAATTCGGGTGGCCATC 120
 QY 115 ACCCGCGCGGCGTGGGAACTAACCGGTGCGGCAATGAGTGTCTCATCCAGGAGAT 174
 Db 121 ACCCGCGCGGCGTGGGAACTAACCGGTGCGGCAATGAGTGTCTCATCCAGGAGAT 180
 QY 175 GCCGAGAGGCGCTCGGCTATCAACGCGGATTTCAAGGCGGAGCGCGCACTGCTC 234
 Db 181 GCCGAGAGGCGCTCGGCTATCAACGCGGATTTCAAGGCGGAGCGCGCACTGCTC 240
 QY 235 GCGACCGCGGACGAGTGTGGGCGGAGTGTATTTATTTCTCAAGTCAAAAGACGAT 294
 Db 241 GCGACCGCGGACGAGTGTGGGCGGAGTGTATTTATTTCTCAAGTCAAAAGACGAT 300
 QY 295 GCGGCGGAATACGCGCGCTGCGACACGCGGAGATCTTGTACGTTTGCATTGGCC 354
 Db 301 GCGGCGGAATACGCGCGCTGCGACACGCGGAGATCTTGTACGTTTGCATTGGCC 360
 QY 355 GCGTCACTGTGTGACGCGATGCTTTGATTCGCGCACACGTCATTTGCTTACGAG 414
 Db 361 GCGTCACTGTGTGACGCGATGCTTTGATTCGCGCACACGTCATTTGCTTACGAG 420
 QY 415 ACCGTCCAGACGCGGAGCGGCGCATACCTGCTTGGCCGATGAGCGAAGTGGCGGT 474
 Db 421 ACCGTCCAGACGCGGAGCGGCGCATACCTGCTTGGCCGATGAGCGAAGTGGCGGT 480
 QY 475 CGACTCGCGCGCCAGGTTGGCTTACACTGTGATCGAAACCAAGGGGCGCGGTGTG 534
 Db 481 CGACTCGCGCGCCAGGTTGGCTTACACTGTGATCGAAACCAAGGGGCGCGGTGTG 540
 QY 535 CTGATGGGGGGGTGGCGCGGCTGCAACGCGGCGGAGTGTGGTGTGGGCGGCGCAC 594
 Db 541 CTGATGGGGGGGTGGCGCGGCTGCAACGCGGCGGAGTGTGGTGTGGGCGGCGCAC 600
 QY 595 GCGGCGTACAAACGCGCGCATGCGCAAGCGGATGGGCGGACCGTTACGGTTCTAGAC 654
 Db 601 GCGGCGTACAAACGCGCGCATGCGCAAGCGGATGGGCGGACCGTTACGGTTCTAGAC 660
 QY 655 ATCAACATGACAAACTTGGCACTCGAGCGGAGTTCTGGCGCGGATCCACACTCGC 714
 Db 661 ATCAACATGACAAACTTGGCACTCGAGCGGAGTTCTGGCGCGGATCCACACTCGC 720
 QY 715 TACTCATCGGCTACGAGCTCGAGGGTGGCGTCAAAAGTGCAGACTGGTGTGGGCC 774

Db 721 TACTCATCGGCTACGAGCTCGAGGGTGGCGTCAAAAGTGCAGACTGGTGTGGGCC 780
 QY 775 GTCCCTGTGCGAGGCGCGCAAGCGCACCAATTAATTCGATTAATTCGCGCATATG 834
 Db 781 GTCCCTGTGCGAGGCGCGCAAGCGCACCAATTAATTCGATTAATTCGCGCATATG 840
 QY 835 AAACAGGTGGGTACTGTGATATAGCCATCGACAGGGGCGCTTTTGAAGGCTCA 894
 Db 841 AAACAGGTGGGTACTGTGATATAGCCATCGACAGGGGCGCTTTTGAAGGCTCA 900
 QY 895 CGACGCGACGCTACGACACCGGAGCGTGGCGGTGCGGACGACGCTGTTTACTGCGT 954
 Db 901 CGACGCGACGCTACGACACCGGAGCGTGGCGGTGCGGACGACGCTGTTTACTGCGT 960
 QY 955 GCGAAGTGGCGCGCTGCGTGGCGGAGAGCTGACCTACGCGTGCACACCGGAGATG 1014
 Db 961 GCGAAGTGGCGCGCTGCGTGGCGGAGAGCTGACCTACGCGTGCACACCGGAGATG 1020
 QY 1015 CCGTATGTGCTGAGCTTGGCGGACATGGCTGGCGGCGGCGGCGGCGGATCGGCA 1074
 Db 1021 CCGTATGTGCTGAGCTTGGCGGACATGGCTGGCGGCGGCGGCGGCGGATCGGCA 1080
 QY 1075 CTAGCCAAAGTCTTTGACGCGACGAGGCGCTTACTGTCCGAACGSGTGGCCACCGAC 1134
 Db 1081 CTAGCCAAAGTCTTTGACGCGACGAGGCGCTTACTGTCCGAACGSGTGGCCACCGAC 1140
 QY 1135 CTGGGGTGGCGCTTCAACCGGCGGCGGAGCGTGTGCGCTGACGCTCGGCGGCTGTTAC 1194
 Db 1141 CTGGGGTGGCGCTTCAACCGGCGGCGGAGCGTGTGCGCTGACGCTCGGCGGCTGTTAC 1200
 QY 1195 GCGGAGCACAGTGGGAGTAAAGGAAAGCATGATGTGGGCC 1236
 Db 1201 GCGGAGCACAGTGGGAGTAAAGGAAAGCATGATGTGGGCC 1242

Search completed: June 22, 2000, 14:57:54
 Job time: 17889 sec

34	42.2	3.4	329	1	V44425	Mycobacterium tube
----	------	-----	-----	---	--------	--------------------

XY 1/3 GCGGAGAGGCTCCGCTATCACC
|||||

1/5 G C C C G A G A G G C C T C G G C T A T C A C C G A C C G G A T T C A A G G C G G C A G G C C G C A C T G G T C 234

```

QY 235 GGCACCCGCGACAGTGTGGGCGGACGCTGATTATTGTCGAAGTCAAGAAACCGATA 294
    |||||||
Db 241 GGCACCGCGCGACAGTGTGGGCGGACGCTGATTATTGTCGAAGTCAAGAAACCGATA 300
QY 295 GCGGCGGAATACGGCCGCTCGGACAGCGGAGATCTTGTACGTTTTCGATTGGCC 354
    |||||||
Db 301 GCGGCGGAATACGGCCGCTCGGACAGCGGAGATCTTGTACGTTTTCGATTGGCC 360
QY 355 GCGTCAGTGTTCAGACGATGCGTGTGATTCGCGCACAGCTCAATGGCTACGAG 414
    |||||||
Db 361 GCGTCAGTGTTCAGACGATGCGTGTGATTCGCGCACAGCTCAATGGCTACGAG 420
QY 415 ACCGTCCAGACCGCGGACGCGGACATCCCTGCTTGCCTCCGATGAGGAAAGTCCGCGT 474
    |||||||
Db 421 ACCGTCCAGACCGCGGACGCGGACATCCCTGCTTGCCTCCGATGAGGAAAGTCCGCGT 480
QY 475 CGACTCGCGCGCGGATGGGCTTACCACTGATGCAACCCAAAGGGGCGCGGTGTG 534
    |||||||
Db 481 CGACTCGCGCGCGGATGGGCTTACCACTGATGCAACCCAAAGGGGCGCGGTGTG 540
QY 535 CTGATGGGCGGGTGGCGGCGGCGTCAACCGGCGAGCTGTGATGCGCGCGGACAC 594
    |||||||
Db 541 CTGATGGGCGGGTGGCGGCGGCGTCAACCGGCGAGCTGTGATGCGCGCGGACAC 600
QY 595 GCGGCGTCAACGAGCGCGCATGCGCAACGCGATGGCGCGACCGTTACGGTTCTAGAC 654
    |||||||
Db 601 GCGGCGTCAACGAGCGCGCATGCGCAACGCGATGGCGCGACCGTTACGGTTCTAGAC 660
QY 655 ATCAACATCGCAAACTCGGCACTGAGCGCGGATCTCGCGCGGATCCACACTCGC 714
    |||||||
Db 661 ATCAACATCGCAAACTCGGCACTGAGCGCGGATCTCGCGCGGATCCACACTCGC 720
QY 715 TACTCATCGGCTACGAGCTGAGGGTCCGTCAAACTGCGGACCTGATTTGGGCGC 774
    |||||||
Db 721 TACTCATCGGCTACGAGCTGAGGGTCCGTCAAACTGCGGACCTGATTTGGGCGC 780
QY 775 GTCTGTGTCGAGCGCGCAAGGACCAAAATAGTCTCGAAATTCATCTTGTGCGCATATG 834
    |||||||
Db 781 GTCTGTGTCGAGCGCGCAAGGACCAAAATAGTCTCGAAATTCATCTTGTGCGCATATG 840
QY 835 AAACGAGTGGGCTGACTGTGGATATAGCATCGACGAGGGGCGGTGTTGGAAGGCTCA 894
    |||||||
Db 841 AAACGAGTGGGCTGACTGTGGATATAGCATCGACGAGGGGCGGTGTTGGAAGGCTCA 900
QY 895 CGACGACACCTACGAGACACCGGACGCTTCCGCGGACGACGCTGTTTACTGCGTG 954
    |||||||
Db 901 CGACGACACCTACGAGACACCGGACGCTTCCGCGGACGACGCTGTTTACTGCGTG 960
QY 955 GCGAATCATGCCGCTCGGTGCGGAAGACGTGACCTACGCGCTGACCAACGCGCATG 1014
    |||||||
Db 961 GCGAATCATGCCGCTCGGTGCGGAAGACGTGACCTACGCGCTGACCAACGCGCATG 1020
QY 1015 CCGATGTGCTGAGCTTGGCGGACCATGGCTGGGCGGGGCTGCGGCTGCAATCCGGCA 1074
    |||||||
Db 1021 CCGATGTGCTGAGCTTGGCGGACCATGGCTGGGCGGGGCTGCGGCTGCAATCCGGCA 1080
QY 1075 CTAGCGCAAGGCTCTTTCAGACGACGAAAGGGGCTTACTGTCCGAAGGGGTGGCCACGAC 1134
    |||||||
Db 1081 CTAGCGCAAGGCTCTTTCAGACGACGAAAGGGGCTTACTGTCCGAAGGGGTGGCCACGAC 1140
QY 1135 CTGGGGGTGCGCTTACCGAGCGCGGACGCTGCGCTGACTCTGCGCGCTGTTAC 1194
    |||||||
Db 1141 CTGGGGGTGCGCTTACCGAGCGCGGACGCTGCGCTGACTCTGCGCGCTGTTAC 1200
QY 1195 GCGGAGCACAGTGGGAGTAAGGGAAGCATGATGGGCGC 1236
    |||||||
Db 1201 GCGGAGCACAGTGGGAGTAAGGGAAGCATGATGGGCGC 1242

```

RESULT 2
 ID V49510 standard; DNA: 1260 BP.
 AC V49510;

```

DT 20-OCT-1998 (first entry)
DE Mycobacterium sp. AladH DNA.
KW Alanine dehydrogenase; AladH; ADH; diagnosis; tuberculosis; pathogen;
KW swimmers disease; vaccine; epidemic; infection; identification; ss.
OS Mycobacterium sp.
PN W09832862-A2.
PD 30-JUL-1998.
PF 29-JAN-1998; E00484.
PR 29-JAN-1997; EP-101339.
PA (FLOH/) FLOHE L.
PI Fione L, Hutter B, Kolik A, Singh M;
DR WPI: 98-427958/35.
PT Nucleic acid encoding alanine dehydrogenase of Mycobacterium marinum
PT - used for diagnosis of tuberculosis and other mycobacterial
PT diseases, also for treatment and prevention, for drug screening and
PT for bio-transformation
PS Disclosure: Page 11; 57pp: German.
CC This sequence encodes an alanine dehydrogenase (ADH) protein isolated
CC from a Mycobacterium sp. This protein is used to diagnose tuberculosis
CC and other mycobacterial infections (including 'swimmers' disease', caused
CC by M. marinum, a fish pathogen) in humans or animals. The protein can
CC also be used for control of epidemics and for vaccination. To screen for
CC agents with anti-mycobacterial activity, and in bio-transformations that
CC are specific for L-alanine. Also Mycobacteria can be identified by
CC analysis of genomic ADH sequences. ADH is an antigen thpt is secreted
CC early during infection.
SQ Sequence 1260 BP; 243 A; 403 C; 389 G; 225 T;

```

Query Match 98.7%; Score 1220; DB 1; Length 1260;
 Best Local Similarity 99.5%; Pred. No. 3,4e-245;
 Matches 1236; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

```

QY 1 ATCTTGAGATTAATGCACTTCTTCACTAGAGCTACATGAGAGGGTAAAC 60
    |||||||
Db 16 ATCTTGAGATTAATGCACTTCTTCACTAGAGCTACATGAGAGGGTAAAC 75
QY 61 ATGCGCTCGGTATTCGACGAGACCAAAACACG-----AATTCGGGTGGCATC 114
    |||||||
Db 76 ATGCGCTCGGTATTCGACGAGACCAAAACAGCAATTCGAGGTGGGCGCATC 135
QY 115 ACCCGGCGCGGCTGCGGACATAACCGTGTGGCCATGAGGTGCTCATCGAGGAGT 174
    |||||||
Db 136 ACCCGGCGCGGCTGCGGACATAACCGTGTGGCCATGAGGTGCTCATCGAGGAGT 195
QY 175 GCGGAGAGGGCTCGGTATCACCGAGCGGATTTCAAGCGGCGGACGCAATGCTG 224
    |||||||
Db 196 GCGGAGAGGGCTCGGTATCACCGAGCGGATTTCAAGCGGCGGACGCAATGCTG 255
QY 235 GGCACCCCGGACGAGTGTGGCGGACGCTGATTTATGCTCAAGGCAAGAACGATA 294
    |||||||
Db 256 GGCACCCCGGACGAGTGTGGCGGACGCTGATTTATGCTCAAGGCAAGAACGATA 315
QY 295 GCGGCGGAATACGGCGCGCTGCGACAGGGGCGAGATCTTGTTCACGTTCTTGAGCC 354
    |||||||
Db 316 GCGGCGGAATACGGCGCGCTGCGACAGGGGCGAGATCTTGTTCACGTTCTTGAGCC 375
QY 355 GCGTCAGTGTTCAGACGATGCGTGTGATTCGCGGACGAGCTCAATGGCTACGAG 414
    |||||||
Db 376 GCGTCAGTGTTCAGACGATGCGTGTGATTCGCGGACGAGCTCAATGGCTACGAG 435
QY 415 ACCGTCCAGACCGCGGACGCGGACGCTACCCCTGCTTCCCGGATGAGCGAAAGTGGCGGT 474
    |||||||
Db 436 ACCGTCCAGACCGCGGACGCGGACGCTACCCCTGCTTCCCGGATGAGCGAAAGTGGCGGT 495
QY 475 CGACTCGCGCGCGGATGGCGCTTACCACTGATGGAACCAAGGGGCGCGGTGTG 534
    |||||||
Db 496 CGACTCGCGCGCGGATGGCGCTTACCACTGATGGAACCAAGGGGCGCGGTGTG 555
QY 535 CTGATGGGCGGGTGGCGGCGTGAACCGGCGGACGCTGTTGATGCGGCGCGGAC 594
    |||||||
Db 556 CTGATGGGCGGGTGGCGGCGTGAACCGGCGGACGCTGTTGATGCGGCGCGGAC 615

```


QY 595 GCGGCTACAGCAGCCGCGATCGCCAGCGATGGGGGCGGACGGTGTACGGTCTAGAC 654
 DB 616 GCGGCTACAGCAGCAGCCGCGATCGCCAGCGATGGGGGCGGACGGTGTACGGTCTAGAC 675
 QY 655 ATCAACATGACAAATCTTGGCACTCGAGCGGAGTCTGGGGCGGATCCACACTCGC 714
 DB 676 ATCAACATGACAAATCTTGGCACTCGAGCGGAGTCTGGGGCGGATCCACACTCGC 735
 QY 715 TACTATCGGCTAGAGCTCGAGGGTGCCTGCAAAAGTGGCCAGCTGTATGGGGCC 774
 DB 736 TACTATCGGCTAGAGCTCGAGGGTGCCTGCAAAAGTGGCCAGCTGTATGGGGCC 795
 QY 775 GTCTGTGTCGAGCGGCGCAAGGACCAATTAATGCTCAATTAATTTCCGGCATATG 834
 DB 796 GTCTGTGTCGAGGGCGCAAGGACCAATTAATGCTCAATTAATTTCCGGCATATG 855
 QY 835 AAACCGAGTGTAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 894
 DB 856 AAACCGAGTGTAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 915
 QY 895 CGACCGAGTGTAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 954
 DB 916 CGACCGAGTGTAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 975
 QY 955 GCGGAGTGTAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1014
 DB 976 GCGGAGTGTAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1035
 QY 1015 CGGATGTGTGAGTGTAGTGTATGATGATGATGATGATGATGATGATGATGATGATG 1074
 DB 1036 CGGATGTGTGAGTGTAGTGTATGATGATGATGATGATGATGATGATGATGATGATG 1095
 QY 1075 CTAGCCAAAGTGTGTTCGACGACGACGAGGAGGCGTACTGTCCGAGGGTGGCCACGAC 1134
 DB 1096 CTAGCCAAAGTGTGTTCGACGACGACGAGGAGGCGTACTGTCCGAGGGTGGCCACGAC 1155
 QY 1135 CTGGGGGTGCGGTTCACGACGACGAGGAGGCGTACTGTCCGAGGGTGGCCACGAC 1194
 DB 1156 CTGGGGGTGCGGTTCACGACGACGAGGAGGCGTACTGTCCGAGGGTGGCCACGAC 1215
 QY 1195 GCGGAGTGTAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1236
 DB 1216 GCGGAGTGTAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1257
 RESULT 3
 V49625
 ID V49625 standard; DNA; 1260 BP.
 AC V49625;
 DT 20-NOV-1998 (first entry)
 DE Mycobacterium tuberculosis Alamine dehydrogenase.
 KM ss: Alamine dehydrogenase: tuberculosis; TB; mycobacteria; L-alanine.
 OS Mycobacterium tuberculosis.
 PN W09836089-42.
 PD 20-AUG-1998.
 PF 29-JAN-1998; E00483.
 PR 29-JAN-1997; EP-101338.
 PA (FLOH/) FLOHE L.
 PI Flohe L, Hutter B, Kolk A, Singh M;
 DR MPI: 98-457123/39.
 PT Diagnosis of, e.g., tuberculosis from alamine dehydrogenase activity
 PT useful for, e.g., for diagnosis, differentiation of strains,
 PT monitoring vaccination and identification of mycobacteria.
 PT inhibitors
 PT Claim 13; Fig 2.3; 55pp; German.
 CC The Mycobacterium tuberculosis alamine dehydrogenase (AlaDH) is in
 CC the production of kits for diagnosing tuberculosis (TB) and other
 CC mycobacterial infections in humans or animals. Kits are used for direct
 CC diagnosis of TB on clinical samples (e.g. body fluids) and can
 CC differentiate between pathogenic and non-virulent strains; e.g. for
 CC identifying pathogens of the M. tuberculosis (M. T.) complex. The kit may
 CC also be used to identify substances that inhibit mycobacteria, for
 CC combating epidemics and for vaccination follow-up. Oligonucleotides

CC derived from AlaDH are used similarly in diagnostic hybridisation tests,
 CC also for culture confirmation of isolated strains and for chromosome
 CC fingerprinting to detect/differentiate between mycobacteria, and for
 CC L-alanine-specific bioassays for differentiation reactions. AlaDH is an early
 CC antigen, present extracellularly after only a few days of growth, making
 CC it an ideal drug target.
 SO Sequence 1260 BP; 243 A; 403 C; 389 G; 225 T;
 Query Match 98.7%; Score 1220; DB 1; Length 1260;
 Best Local Similarity 99.5%; Pred. No. 3.4e-245;
 Matches 1236; Conservative 0; Mismatches 0; Indels 6; Gaps 1;
 QY 1 ATCTTGACAGATTAATGACACTTCTTACACTGAGCGTACGATGACAGGGTAAATC 60
 DB 16 ATCTTGACAGATTAATGACACTTCTTACACTGAGCGTACGATGACAGGGTAAATC 75
 QY 61 ATGGCGGTGGTATTCGCGAGACCAAAACAAACG-----AATTGCGGGTGGCCATC 114
 DB 76 ATGGCGGTGGTATTCGCGAGACCAAAACAAACGATTCGATTCGAGGGGTAAATC 135
 QY 115 ACCCGGCGCGCGTGGCGGACCACTACCGCTGTCGATGAGTGTCTATCAGGCAAGT 174
 DB 136 ACCCGGCGCGCGTGGCGGACCACTACCGCTGTCGATGAGTGTCTATCAGGCAAGT 195
 QY 175 GCGGAGAGGCTCGGCTATACAGCAGCGGATTTCAAGGGCGGAGCGGCACTGCTC 234
 DB 196 GCGGAGAGGCTCGGCTATACAGCAGCGGATTTCAAGGGCGGAGCGGCACTGCTC 255
 QY 235 GGCACCGCGGACGAGTGTGGCGGAGCGTATGATGATGATGATGATGATGATGATGATG 294
 DB 256 GGCACCGCGGACGAGTGTGGCGGAGCGTATGATGATGATGATGATGATGATGATGATG 315
 QY 295 GCGGCGGATTAACGCGCGCTGCGACAGCGGAGATCTTGTACGCTTTCGATTTGGCC 354
 DB 316 GCGGCGGATTAACGCGCGCTGCGACAGCGGAGATCTTGTACGCTTTCGATTTGGCC 375
 QY 355 GCGTACAGTGTTCGACCG "GCGTGTGATTCGCGGACCA TCAATTCCTACGAG 414
 DB 376 GCGTACAGTGTTCGACCG "GCGTGTGATTCGCGGACCA TCAATTCCTACGAG 435
 QY 415 ACCGTCACAGCGCGGAGCGGACGACGCGGATGCGGATGCGGATGCGGATGCGGAT 474
 DB 436 ACCGTCACAGCGCGGAGCGGACGACGCGGATGCGGATGCGGATGCGGATGCGGAT 495
 QY 475 CGACTGCGCGCGGAGGCTTACGACGATGATGACCAAGGAGGCGCGGATG 534
 DB 496 CGACTGCGCGCGGAGGCTTACGACGATGATGACCAAGGAGGCGCGGATG 555
 QY 535 CTGATGGGCGGCGGCGCGGCTTCGACCGCGGAGCGGATGCGGATGCGGATGCGG 594
 DB 556 CTGATGGGCGGCGGCGGCGGCTTCGACCGCGGAGCGGATGCGGATGCGGATGCGG 615
 QY 595 GCGGCTACAGGCGCGGCGGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGG 654
 DB 616 GCGGCTACAGGCGCGGCGGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGG 675
 QY 655 ATCAACATGACAAATCTTGGCACTCGAGCGGAGTCTGGGGCGGATCCACACTCGC 714
 DB 676 ATCAACATGACAAATCTTGGCACTCGAGCGGAGTCTGGGGCGGATCCACACTCGC 735
 QY 715 TACTATCGGCTAGAGCTCGAGGGTGCCTGCAAAAGTGGCCAGCTGTATGGGGCC 774
 DB 736 TACTATCGGCTAGAGCTCGAGGGTGCCTGCAAAAGTGGCCAGCTGTATGGGGCC 795
 QY 775 GTCTGTGTCGAGCGGCGCAAGGACCAATTAATGCTCAATTAATTTCCGGCATATG 834
 DB 796 GTCTGTGTCGAGCGGCGCAAGGACCAATTAATGCTCAATTAATTTCCGGCATATG 855
 QY 835 AAACCGAGTGTAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 894
 DB 856 AAACCGAGTGTAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 915

Qy	895	CGACCGACACACCTTAACGACACACCCGAGGTTGGCCGCTGACAGACACGGTTATTACGCGTG	954
Db	916	CGACCGACACACTTAACGACACACCCGAGGTTGGCCGCTGACAGACACGGTTATTACGCGTG	975
Qy	955	GCAGAACATGCCCGCGCTCGGAGCCGAAGACGCTGCACCTACGCGCTGCACCAACGCGACGATG	1014
Db	976	GCAGAACATGCCCGCGCTCGGAGCCGAAGACGCTGCACCTACGCGCTGCACCAACGCGACGATG	1035
Qy	1015	CCGATATGTGCTCGAGCTTGGCCGACATAGCTGGCGGGCGCGTGGCCGCTCGAATCGGCA	1074
Db	1036	CCGATATGTGCTCGAGCTTGGCCGACATAGCTGGCGGGCGCGTGGCCGCTCGAATCGGCA	1095
Qy	1075	CTAGCCAAAGGCTCTTTGACGACGACGAAGGGGCTTACTGTCCGAACGGGTGGCCACACGAC	1134
Db	1096	CTAGCCAAAGGCTCTTTGACGACGACGAAGGGGCTTACTGTCCGAACGGGTGGCCACACGAC	1155
Qy	1135	CTGGGGGTCGCCGTTCCACCGAGCCCGCCAGGTCCTAGCCGATCTCGGCGCGCTGTTAC	1194
Db	1156	CTGGGGGTCGCCGTTCCACCGAGCCCGCCAGGTCCTAGCCGATCTCGGCGCGCTGTTAC	1215
Qy	1195	GCCGAGCACACGTCGGGAGTAAGGAAGCATATGTCGGC	1236
Db	1216	GCCGAGCACACGTCGGGAGTAAGGAAGCATATGTCGGC ; 1257	

RESULT	4
V49511	
ID	V49511 standard; DNA; 682 BP.
AC	V49511.
DT	20-OCT-1998 (first entry)
DE	Mycobacterium marinum Mar3 DNA.
KM	Alanine dehydrogenase; AlaDH; ADH; diagnosis; tuberculosis; pathogen.
KW	Swimmers disease; vaccine; epidemic; infection; identification; ss.
OS	Mycobacterium marinum.
FH	Location/Qualifiers
Key	1..682
CDS	/*tag= a
FT	/codon_start= 3
FT	/product= "Mar3"
FT	/note= "Alanine dehydrogenase"
FT	

PN WO9832862-A2.
PD 30-JUL-1998.
PF 29-JAN-1998: E00484.
PR 29-JAN-1997: EP-101339.
PA (FLOH/) FLOHE L. Singh M:
PI Fione L. Rutter B. Kolik A.
DR WPI: 98-427958/36.
PT P-PDB: W64A81.
PT Nucleic acid encoding alanine dehydrogenase of *Mycobacterium marinum*
PT - used for diagnosis of tuberculosis and other mycobacterial
PT diseases, also for treatment and prevention, for drug screening and
PT for bio-transformation
PS Claim 1, Page 34-35; 57pp; German.
CC This sequence encodes an alanine dehydrogenase (ADH) protein, Mar3
CC isolated from *Mycobacterium marinum*. This protein is used to diagnose
CC tuberculosis and other mycobacterial infections (including 'swimmers'
CC disease', caused by *M. marinum*, a fish pathogen) in humans or animals.
CC The protein can also be used for control of epidemics and for
CC vaccination, to screen for agents with anti-mycobacterial activity, and
CC in bio-transformations that are specific for L-alanine. Also mycobacteria
CC can be identified by analysis of genomic ADH sequences. ADH is an antigen
CC that is secreted early during infection.
SQ Sequence 682 Bp; 105 A; 254 C; 225 G; 98 T;

	Query Match	37.8%	Score	467.6	DB 1	Length	682:			
	Best Local Similarity	80.4%	Pred. No.	5.1e-89						
	Matches	568:	Conservative	0:	Mismatches	134:	Indels	0:	Gaps	0:
QY	98	AATTCGGGTGCCATACCCGCGCGCGTGCAGACTAACCCCTGTGGCCATTAG	157							
Db	1	AATTCGGGTAGATCACCCGCGCGCGTGCAGACTTGACAAGCGCGCGCAGAG	60							

Oy	158	TGTCATCTCAGCAGAGGTGTCGGAGAGAGGGCTTCGGCTATACACAGACGGGAGATTCAAGCCG	217
Db	61	TGCTGATCTCAGGCGCGGTGTCGGAGAGAGGCTTCGGCATCTCCGACGCCGACTTCAGGCCG	120
Oy	218	CAGGCGCGCAACTGTTGCGGACCGCGCCAGCAGGTGTGTGGGCGGACGCTGATTTATGTCTCA	277
Db	121	CCGGTGCCACACTGATCATGACACCGCCGACAGGTGTGTGGGCGGATGTGGAGACTCTGCTCA	180
Oy	278	AGGTCGAAGAACCGATAGCGGGCGGAATTCCTCTGCTGCGACACAGGGGCAATCTTTCTCA	337
Db	181	AGGTCGAAGAACCGATGAGTCGAGTAGCGGCGGTGCGCCGCGGGCGCAACCTTCTTA	240
Oy	338	CGTCTTCTCATTTGGCCGCGTCACAGTCTCTTGACCCGATGCGTGTGTGGATTGCTCCGAGCCA	397
Db	241	CCTACTCTCAGCTGGCCGCTCGCGGCCCTCTGACACCGATGCGCTGTGTAATCGGAGCCA	300
Oy	398	CGTCAATTGCTTACGAGACCGTCCAGACCGCGCGAGCGGCACTACCCCTGCTTGGCCCCA	457
Db	301	CGTCAATGCGCTTACGAGACCGGTGAGACCCCGACCGCGGATTTGCGGCTGCTTGGCCCCA	360
Oy	458	TGAGGGAAGTGGCGGCTGCGACTGCGCGCCGAGGTGTGGCGTTACACACTGATGCGAACC	517
Db	361	TGAGGGAAGTGGCGGCGGCGGCTGTCCGCCCAAGTGTGGGCGCTTACCACTTATGCGACCC	420
Oy	518	AAGGGGCGCGGCGGTGTGCTGATGGGCGGGGTGCCCGGCGTTCGAACCGCGCGAGCTGTGG	577
Db	421	ACGGCGGTGCGGCGGTGTGATGGGCGGCGTCCCGGCGTCAAGCCTTGCAGCGTGTGG	480
Oy	578	TGATGGGGCGCGGACCGCGCGGCTACACAACGACGCCGCTATGCGCAAGGATGGGCGCA	637
Db	481	TGATGGGGCGGGGACCGCGCGGATACCAACGCGCGCGCGCTGCGCAAGGATGGGCGCA	540
Oy	638	CCGTTACGGTCTAGACATCAACATTCACAACAACTTGGGAACTGAGCGCGAGATTGTGCG	697
Db	541	TGGTCAACGCTCGGTGATTCAAATCAACAACGCTCCCGCAGATGACAGCGGATTTGGCG	600
Oy	698	GCCGATCCACACTGCTACTCATCGGCTACGAGGCTGAGAGGTGCCGTAAACGTGGCG	757
Db	601	GTCGCGTCCCGAGCCGCTACTGCTGACCCCTGCACTCTGAGGATGCGGCGAGTCCACGCCG	660
Oy	758	ACCTGGTATGGGGCGCT f 779	
Db	661	ACATGATGTCGGGCGGCTCT 682	

RESULT 5
ID N91423 standard; DNA: 1125 bp.
AC N91423.
DT 01-FEB-1991 (first entry)
DE Sequence of heat-resistant alanine dehydrogenase (AH) gene with mol. wt. 2 Md or less
KW Enzyme: ds.
OS *Bacillus stearothermophilus* IFO 12550.
PN J01043194-A.
PD 15-FEB-1989.
PF 10-AUG-1987; 200524.
PR 10-AUG-1987; JP-200524.
PA (NIRA) Unilika KK.
DR MPI: 89-096096/13.
PT Recombinant plasmid for transforming *Escherichia coli* -
PI obtc. by connecting heat resistant alanine dehydrogenase gene to vector plasmid
PS Disclosure: Fig 2, p693; 9pp; Japanese.
CC A recombinant plasmid confg. heat-resistant alanine dehydrogenase (AH) gene with mol. wt. 2 or less Md is claimed. Cells transformed with the CC vector produce high levels of heat-resistant AH. Transformed *E. coli* is useful as a clinical inspection reagent.
SQ Sequence 1125 bp; 246 A; 295 C; 367 G; 217 T;

Query Match	22.4%;	Score 277.2;	DB 1;	Length 1125;
Best Local Similarity	56.2%;	Pred. NO. 1.8e-49;		g


```

QY 361 CGTGTTCACCGATGCTGTGGATTCGGGACACGATTCATTCGACGACCCGTC 420
DB 21414 CCAGAAATTTACGATGCTATGTT-----AACACCAAAAACACTGAACCTGTT 21367
QY 421 CAGACCGCGACGCGGACGATACCCCTGCTGCCCCGATGAGCGAAGTCGCCGGCGACTC 480
DB 21366 CGTGCACATCAAGGACGACATCCGCTCTGCTTCTATGAGTGAGTGAGGCGATG 21307
QY 481 GCGGCCGAGGTGGCGCTTACACCTGATCGAACCCAGGGGGCCGCGGTGCTGATG 540
DB 21306 GCTGTTCATTAATTCGAGAGTCACTTCTACTAGAACAGCTGTGCTGCTGCTTCTACTT 21247
QY 541 GCGGGGTGCGCGCGCTGACACCGCGGAGCTGTGTATTCGCGCGCGCGCGCGCG 600
DB 21246 GGTGTGTACAGAGTGTTCACAAAAGAAAGTAACTATCGGTGTGTGTGTGTGTGT 21187
QY 601 TACAAAGCAGCGCGCATCGGCAAGGATGGGCGCGACCGCTTACGCTTACAGATCAAC 660
DB 21186 ACACATGCTGCGCGCATCGGCGCGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 21127
QY 661 ATGCAAACTTGGCACTCGACCGCGGAGTTCGCGCGGATCCACACTCGCTACTCA 720
DB 21126 TCCAAAGGCTCTCTAGTTCTAGAAAGATGCTTGGAGTCA/ATTCAAACTTATGTCT 21067
QY 721 TCGGCTTACGAGCTCGAGGCGCGCTCAACGCGCGACCTGCTGATGGCGCGCTCTG 780
DB 21066 AATTCAATTAACATTAAGCAAGTGTAGAGATGCTATGTTGATTTGAGGCGCTTCTC 21007
QY 781 GTGCCAGGCGCCCAAGCAGCCCAATTAATGTCTGCAATTCAGTTCGCGCATATGAACA 840
DB 21006 ATCCCTGTGTCAAAAGCAGCGAATGTGTGACAGATGATGATGCAAAAGTCCGCTCA 20947
QY 841 GGTGCGTACTGTGTGATTAATGACATGACACGCGCGCTGTTCGA---AGGCTACGA 897
DB 20946 GGCTCTGTA-TGCTTGAAGTGTGCTGTTGACCAAGGTGCGCTTATACGAACAGCTGACGT 20888
QY 898 CCGACCACTTACGACGACCGGAGCTTGGCGGTGCGGACAGACGCTTCTGCTGCGGCG 957
DB 20887 GTTGCACACGACGATGAAACCGCTTATGAAAAACACGCTGTCTCCACTATGCCGTTGCC 20828
QY 958 AACATGCGCGCGCTGCTGTCGGAAGACGCTGACGCTGACGCAACGCGGACGATGCCG 1017
DB 20827 AATATCCCTGTGTGTGCTGCTGCGCACTTCAACCATGCCCTTAACCAATGTACACTTCTCT 20768
QY 1018 TATGTGCTCGAGCTTGGCGGACCATGCTT 1045
DB 20767 TATATCGAAGCTTGTGCTGCGCAAGGAT 20740

RESULT 7
V74442/c
ID V74442 standard; DNA; 9280 BP.
AC V74442;
DE 16-MAR-1999 (first entry)
KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KM skin infection; surgical wound infection; scalded skin syndrome; ds.
OS Staphylococcus aureus.
FH Key location/Qualifiers
misc_feature 841..900
FT /tag- a
FT /note- "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT misc_feature 2641..2700
FT /tag- b
FT /note- "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"

```

```

FT misc_feature 4441..4500
FT /tag- c
FT /note- "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT misc_feature 6241..6300
FT /tag- d
FT /note- "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT misc_feature 8041..8100
FT /tag- e
FT /note- "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT EP-786519-A2.
FT 30-JUL-1997. 100117.
FT 07-JAN-1997. 100117.
FT 05-JAN-1996; US-009861.
FT (HUMA-) HUMAN GENOME SCI INC.
FT Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA,
FT Rosen CA;
FT WPI; 97-374922/35.
FT Polynucleotide(s) and proteins derived from Staphylococcus aureus
FT stored on computer readable medium and used in the production of
FT anti-S.aureus vaccines
FT Claim 1: Page 710-715: 3271pp; English.
FT This sequence represents one of 5191 Staphylococcus aureus DNA sequences
FT of the invention. The DNA sequences are recorded on a computer readable
FT medium, preferably selected from a floppy or hard disk, random access
FT memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
FT the S.aureus DNA sequences allows putative functions to be assigned so
FT that protein-encoding or regulatory regions of commercial, therapeutic or
FT industrial importance can be obtained. Specifically, sequences which are
FT likely to encode antigens have been identified and these polypeptides can
FT be used in a vaccine composition against S.aureus infection. The
FT polypeptides can also be used in a kit for the immunodetection of
FT S.aureus in a sample. S.aureus is implicated in numerous human diseases,
FT including cellulitis, eyelid infections, food poisoning, osteomyelitis,
FT skin and surgical wound infections, scalded skin syndrome, toxic shock
FT syndrome, etc. Organisms transformed with the DNA sequences can be used
FT for recombinant production of the polypeptides. The new DNA sequences
FT (and their fragments) are useful as primers or probes for isolating
FT CC homologues of any of the S.aureus DNA sequences contained on the
FT CC computer readable medium.
FT Sequence 9280 BP; 2958 A; 1649 C; 1309 G; 3060 T;

Query Match 14.4%; Score 178.4; DB 1; Length 9280;
Best Local Similarity 50.6%; Pred. No. 6,7e-29;
Matches 484; Conservative 0; Mismatches 466; Indels 6; Gaps 2;

QY 151 CAGAGAGTGTCTATCCAGCAGGTGCGGAGAGGCTCGGCTATCCAGCAGGATTTTC 210
DB 9280 CATACTGTTTAACTGGAACAAACAAATGCGGGTTCAGGATCATTTCTTAATGATGATTAAC 9221
QY 211 AAGCGCGACGCGCGGCAACGATGCTGCGGACCGCGGACGAGTGTGGCGGACGCTGATTA 270
DB 9220 AAAGGAAGCAGGTGCTGAGATGTTGCTGGAACAAACAAAGTTTGG---GATGTGATATG 9164
QY 271 TTGCTCAAGGTCAAAAGACGATAGCGGGAATAGCGCGCTGCGACAGCGGCAATC 330
DB 9163 GTATATTAAGTTAAAGAACACATGATATGATATTCATATTTAAAGAGGCTGTGA 9104
QY 331 TTGTTACGTTCTTTCATTTGGCGGCTACGCTGCTTGCACCGCATCGTTGTGATTC 390
DB 9103 TTATTTCACTTATTTCTTATTCGCAATGAAAGAAATTAACACACTTGTATGATAGA 9044
QY 391 GGCACCGCTGCAATGCTGCTGAGACCGCTCACAACGCGGCGGCGCATACCCCGCTT 450

```

Db 9043 AAGTAATAGTATTGATATGATGACTGTGCACTTACAGACCACTTTTACCATTTGTA 8984
 QY 451 GCCCGATAGAGGAGTCCCGGTGACACCGCCGAGTTGGCGCTTACACACCGATG 510
 Db 8983 TCACCAATAGTAGAGTACAGAGAAATGTACGCTCAAGTTGGCCAGATTCTACAA 8924
 QY 511 CCACCAAGAGGAGGCGCGGTGTGTGATGGCGGGGTGCGCGGTGCGACCGCGAC 570
 Db 8923 AACTTAAATGAGTGTATGGGAATTTCTACAGTGGTGTGCCAGAGTACTTAAGGTAA 8864
 QY 571 GTGTGTGTATGGCGCGCGGACCGCGCTTACACGACCGCGCATCGCCAAAGCATG 630
 Db 8863 GAACTAATATGAGT 8804
 QY 631 GCGGCGACCGTACGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 690
 Db 8803 GGTGCAATGTATGAGT 8744
 QY 691 TTCTGCGCGCGGATCCACATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 750
 Db 8743 TTCTGCGCGCGGATCCACATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8684
 QY 751 CGTCCGCGACCTGT 810
 Db 8683 CAAAGTATTTAGTATAGT 8624
 QY 811 TCGAATTCATGT 870
 Db 8623 ACAGAGACATGTATTAACAAATGAAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8564
 QY 871 CAGGCGCGCTGT 927
 Db 8563 CAAAGCGGATTTTGAACAACTGTATTAATTAAGACACATGATGATCTCATATATAT 8504
 QY 928 GTGCACGACACGCTGT 987
 Db 8503 AAGCATGT 8444
 QY 988 ACCTACGCGCTGACCAAGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1047
 Db 8443 ACCTAGCTTTAAATATCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8384
 QY 1048 CGGCGCGCTGCGCGT 1103
 Db 8383 AGAGAGCATTTAAATCAATCAACATATATCATTTAGTTAAATCTTAAACAAAG 8328
 RESULT 8
 T67971
 ID T67971 standard; DNA; 1074 BP.
 AC T67971:
 DE 15-JUL-1997 (first entry)
 KW H. pylori membrane protein ORF 05cp20518orf1.
 KW Vaccine; prevention; treatment; infection; identification;
 KW binding compound; bacterium; life cycle; activator; bacteria;
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis;
 KW membrane; amino acid; metabolism; ds.
 OS Helicobacter pylori.
 FH key location/Qualifiers
 FH cds 1..1074
 FT /tag= a
 FT /note= "no stop codon given"
 PN W09640893-A1.
 PD 19-DEC-1996.
 PE 06-JUN-1996: U09122.
 PR 07-JUN-1995: US-487032.
 PR 01-APR-1996: US-630403.
 PA (ASTR) ASTRA AB.
 PI Berghindh OT, Smith D, Mellgaard BL.
 DR P-PSDB: W20718.
 PT Helicobacter pylori nucleic acid sequences and related
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori

PT infection, and to detect Helicobacter
 PS Claim 27: Page 819: 1481bp: English.
 CC The present sequence encodes a Helicobacter pylori membrane
 CC protein likely to contain four membrane spanning regions.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds,
 CC useful as potential H. pylori life cycle activators or inhibitors.
 CC The genomic sequence of H. pylori (ATCC 55679) was determined from
 CC overlapping contigs generated by mechanically shearing the
 CC bacterial DNA. The sequences were analysed for ORF of at least 180
 CC nucleotides, and the predicted coding regions defined by computer
 CC evaluation. To identify likely H. pylori antigens for vaccine
 CC development, the amino acid sequences predicted from various ORF
 CC were analysed for significant homology to other known or expected
 CC membrane proteins. Having identified and determined the sequences
 CC of interest, particular regions can be isolated from H. pylori by
 CC PCR amplification for recombinant polypeptide production, e.g. in
 CC E. coli hosts.
 SQ Sequence 1074 BP; 302 A; 189 C; 297 G; 286 T;

Query Match 7.7%; Score 94.8; DB 1; Length 1074;
 Best Local Similarity 46.1%; Pred. No. 1.3e-11;
 Matches 370; Conservative 0; Mismatches 417; Indels 15; Gaps 1;

QY 265 GATTATTGCTCAGGTCAGCAAGCAACGATAGGCGCGGAATGCGCGCTGACACGCGG 324
 Db 133 GATTGTGTGTGTAATGCAAGAGCTTTAGAGCATGTAATACCTTGTCTCAAGAAAA 192
 QY 325 CAGATCTTGTTCACGCTTCTGTGATTTGGCCGCGTACAGTGTGTGACGATGCTGTGTG 384
 Db 193 GCGACCTGTGTATGTTATTTGATTTAGCGTATCAAAAAAGCTTGTGAATGTTTAT 252
 QY 385 GATTCCGCGACGATCAATTTGCTTACAGACGCTCCAGACCGCGGACGCACTACCC 444
 Db 253 AATAAAAAATATCACTTCTATTGTGCACTGAACCATTTGCGGCGCTAAAGAGACTACC 312
 QY 445 CTGCTGCCCGGATGAGGAAGTGGCGCGTGCACGCGCGCCAGCTGGC----- 496
 Db 313 ATTTTAGCGCTATAGCGGT 372
 QY 496 -----GCTTACACCTGATCGCAACCAAGGGGCGCGCGTGTGTGTGTGTGTGTGTGT 549
 Db 373 TTACTGGCTTTAGACATGTTAAAGGTTTATAGGTAAAGGGGTATCTGAGGGGTTTG 432
 QY 550 CCGGCGCTGCAACCGCGGACCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 609
 Db 433 TCGGGTGGCGCAAGGGGCTAAATTCGTGTAATTGGAGGCGGTGTGTGTGTGTGTGTGT 492
 QY 610 GCCCGATGCGCAACGCGATGGGCGGCGGCTTACGCTTGTGTGTGTGTGTGTGTGTGTGT 669
 Db 493 GCGAAAGCTTTAAGCAATAGGGGCTTAAGTAACGATTTTGAATTTAGCTACCTTAA 552
 QY 670 CTTCGCAACCTGACGCGAGT 729
 Db 553 TTACAAACACACCTTATATATTTGTATGTTAAGAGTCTTAAGGTGAAGTAAGCAACC 612
 QY 730 GAGCTCGAGGGTCCGCTCAACAGCTCCGACCTGTGTGTGTGTGTGTGTGTGTGTGTGT 789
 Db 613 AATATCATTTCAAGCCCTTAAACGGGCGGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 672
 QY 790 GCCAAGGCAACCAATTTGT 849
 Db 673 AGCCAAACCCGTAAGTATCTTAAAGGCTTTTAAATACATGACACAGCAAGGGTTA 732
 QY 850 CTGTGTGATATAGCATGACAGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 909
 Db 733 GTCATTTGATGT 792
 QY 910 GACCAACCGAGCTTGGCGGTGACGACACGCGTTTACTGTGTGTGTGTGTGTGTGTGTGTGT 969
 Db 793 TCTAACCGGCTGTATGTGAGAAAGTTGTTGCAATATATGCGGTGCCACATGCGCAGGG 852

